

Universidade de Lisboa
Instituto de Geografia e Ordenamento do Território



**Sistema Espacial de Apoio à Decisão em Saúde
Doação e Transplantação em Portugal**

Paula Cristina Gonçalves Barroso

Relatório de Estágio orientado
pelo Professor Doutor Fernando Jorge Pedro da Silva Pinto da Rocha
e pelo Dr. Celso Miguel Moreira Marques

Mestrado em Sistemas de Informação Geográfica e Modelação
Territorial Aplicados ao Ordenamento

2019

Universidade de Lisboa
Instituto de Geografia e Ordenamento do Território



Sistema Espacial de Apoio à Decisão em Saúde
Doação e Transplantação em Portugal

Paula Cristina Gonçalves Barroso

Relatório de Estágio orientado
pelo Professor Doutor Fernando Jorge Pedro da Silva Pinto da Rocha e
pelo Dr. Celso Miguel Moreira Marques

Júri:

Presidente: Professor Doutor Nuno Manuel Sessarego Marques da Costa do
Instituto de Geografia e Ordenamento do Território da Universidade de
Lisboa;

Vogais:

- Técnica Superior de Sistemas de Informação Geográfica Rita da Mata Graça Roquette e Silva do Instituto Nacional de Saúde Doutor Ricardo Jorge - Departamento de Epidemiologia - Ministério da Saúde;
- Professor Doutor Fernando Jorge Pedro da Silva Pinto da Rocha do Instituto de Geografia e Ordenamento do Território da Universidade de Lisboa.

2019

Resumo

Portugal ocupa hoje um lugar de destaque em relação à doação mundial de órgãos. A transplantação é um tratamento eficaz a salvar vidas, por vezes a única esperança para quem aguarda um transplante. Contudo têm-se disponibilizado poucas ferramentas que possibilitem o estudo geográfico destes fenómenos, em que o apoio das tecnologias de informação geográfica se pode revelar útil.

O presente relatório de estágio assentou em quatro pontos fulcrais de atuação, sendo eles: i) a deteção de potencial de doação nos municípios de Portugal; ii) a identificação de municípios onde, face às patologias associadas aos doentes transplantados, há a possibilidade de outros afetados pela mesma patologia virem a contribuir para o número de transplantados; iii) a identificação de fluxos de acesso de dadores efetivos; e iv) a identificação fluxos de acesso de doentes transplantados. Estrategicamente, o IPST pretende implementar “a utilização das Tecnologias de Informação e Comunicação (TIC) para potenciar a mudança e a modernização administrativa através da integração dos sistemas de informação (SI), aumentando a fiabilidade e interoperabilidade” onde este projeto se enquadra.

Primordialmente procedeu-se a uma recolha de dados, relativamente aos dadores de órgãos *post mortem* e doentes transplantados, assim como de informação complementar com o intuito de criação e estruturação de uma base de dados. Dada a informatização, geocodificação e georreferenciação da informação, foi possível analisar o comportamento dos dadores e doentes transplantados espacialmente e ao nível dos hospitais do Serviço Nacional de Saúde.

Através de métodos estatísticos, foram realizadas análises com vista a identificar padrões espaciais e constatar se a distribuição espacial do número de dadores e doentes transplantados é aleatória ou aglomerada (em *clusters*). Com a análise de potencial, é possível perspetivar se alguns aspetos populacionais influenciam o número de dadores, ou se existe alguma especificidade na associação de patologias a quem necessita de transplante, e se sim, em que locais.

Palavras-chave: Análise espacial; Doação; Saúde; SIG; Transplantação.

Abstract

Nowadays, Portugal has an important place worldwide in what organ donation is concerned. Transplantation is an effective life-saving treatment, sometimes the only hope for those who are looking for a transplant. However, few tools have been provided to allow the geographic study of these phenomena and thereon technologies of geographic information can be useful.

The main purpose of this internship and study relies on four key points of action: i) Potential detection for donations in the Portuguese cities ii) identification of the municipalities where the number of transplanted people is likely to grow given the pathologies associated with their health conditions iii) identification of the flow of access of organs donors; and iv) identification of the flows of access of transplanted sick patients. Strategically, the IPST wants to implement “the usage of Communications and Information Technology to potentiate the change and the administrative modernization through integration of information systems, increasing the reliability and interoperability” where this project fits.

Primarily, data were collected on organ donors post mortem and on transplanted patients, in order to create a database. Given the computerization, geocoding and georeferencing of information, it was possible to analyze the behavior of donors and transplanted patients spatially and in hospitals of the National Health Service.

Through statistical methods, density analyses were performed in order to prospect if the spatial distribution of transplanted donors and patients is random or agglomerated (in clusters). Through an analysis of potential, it is possible to prospect if the population aspects could have influence on the number of donors, or if exist some specificity in the association of pathologies related to who needs the transplant, and if so, in which locals this happens.

Keywords: Donation; Health; GIS; Spatial analysis; Transplantation.

Índice geral

Capítulo 1. Introdução	1
1.1. Doação e Transplantação em contexto Mundial	2
1.2. Transplantação de Órgãos em Portugal	5
1.3. Doação de Órgãos de dador falecido.....	6
1.4. Instituto Português do Sangue e da Transplantação, IP	9
1.5. A Coordenação Nacional da Transplantação	10
1.6. Área de Estudo	13
Capítulo 2. Geografia e SIG	15
2.1. Geografia da Saúde	16
2.2. Sistemas de Informação Geográfica	17
2.3. Relevância dos SIG na Saúde Pública	18
2.4. Sistema Espacial de Apoio à Decisão	22
2.4.1. Componentes e tipologias do SEAD	23
2.4.2. WebSEAD	25
2.5. Operacionalização do SEAD	25
2.5.1. Obtenção da Informação	26
2.5.2. Construção de uma base de dados em Microsoft Access.....	27
2.5.3. Geocodificação e Georreferenciação	30
Capítulo 3. Métodos de análise e classificação de dados	33
3.1. Índice de Moran	35
3.2. Análise de Hot Spot (Getis-Ord Gi*)	37
3.3. Regressão Linear Múltipla	39
3.3.1. Validação	41
3.3.2. Vantagens e desvantagens.....	41
3.4. Árvores de classificação e regressão	42

3.4.1. Regras de divisão (medidas de impureza dos nós)	42
3.4.2. Cálculo da exatidão na regressão.....	45
3.4.3. Vantagens e desvantagens.....	45
3.5. Redes Neurais Artificiais	47
3.5.1. Perceptrão multi-camada (MLP)	48
3.5.2. O Algoritmo de Retropropagação (avaliação do erro).....	50
3.5.3. Vantagens e desvantagens.....	51
Capítulo 4. Exploração de Dados e Análise Espacial	53
4.1. Dadores de órgãos efetivos	53
4.2. Doentes transplantados.....	69
Capítulo 5. Análise dos Resultados.....	79
5.1. Dadores efetivos	79
5.1.1. Índice Local de Moran	79
5.1.2. Regressão Linear Múltipla.....	80
5.1.3. Árvores de classificação e regressão.....	84
5.1.4. Redes Neurais Artificiais	86
5.2. Doentes Transplantados	89
5.2.1. Índice Global de Moran.....	89
5.2.2. Análise de Hot Spot (Getis-Ord G_i^*).....	90
5.2.3. Regressão Linear Múltipla.....	95
5.2.4. Árvores de classificação e regressão.....	98
5.2.5. Redes Neurais Artificiais	100
Capítulo 6. Considerações Finais	103
Referências Bibliográficas.....	107

Índice de figuras

Figura 1.1. Doação de órgãos em contexto europeu por milhão de habitantes no ano de 2017 (fonte: <i>Newsletter Transplant</i> , 2018)	4
Figura 1.2. Estrutura orgânica da Coordenação Nacional da Transplantação.	11
Figura 1.3. Unidades de transplantação e hospitais com Unidade de Cuidados Intensivos (UCI).	13
Figura 2.1. Componentes de um SIG.	18
Figura 2.2. Casos de cólera durante a epidemia em Londres data de 1854 (Gilbert, 1958).	19
Figura 2.3. Desenvolvimento do SEAD (adaptado de Sugumaran e Sugumaran, 2005:4).	23
Figura 2.4. Componentes de um SEAD. (adaptado de Maczewski, 1997:4).	24
Figura 3.1. Distribuição normal dos <i>p-values</i> e dos <i>z-scores</i> (ArcGis help).	38
Figura 3.2. Esquema de uma MLP completamente conectada. Note-se que o nível escondido e o de saída possuem bias.	48
Figura 4.1. Número total de dadores efetivos nos hospitais com neurocirurgia.....	55
Figura 4.2. Número total de dadores efetivos nos hospitais sem neurocirurgia.....	56
Figura 4.3. Dadores efetivos por camas de cuidados intensivos e intermédios, nos hospitais com neurocirurgia ao nível do município	58
Figura 4.4. Dadores efetivos por camas de cuidados intensivos e intermédios, nos hospitais sem neurocirurgia ao nível do município.....	59
Figura 4.5. Número de dadores efetivos por município no ano de 2016	60
Figura 4.6. Número de dadores efetivos por município no ano de 2017.	61
Figura 4.7. Número de órgãos colhidos e respetiva taxa de utilização nos hospitais pertencentes à NUTS II no ano de 2017.....	62
Figura 4.8. Número de órgãos colhidos e respetiva taxa de reprovação de órgãos nos hospitais pertencentes às NUTS II no ano de 2017.	62
Figura 4.9. Idade média dos dadores efetivos por NUTS II.	63
Figura 4.10. População residente entre os 45 e os 65 anos de idade por município. ...	64
Figura 4.11. Dadores efetivos por género nos hospitais do SNS no ano de 2016.	64
Figura 4.12. Dadores efetivos por género nos hospitais do SNS no ano de 2017.	65
Figura 4.13. Dadores efetivos por género.	65

Figura 4.14. Distribuição dos dadores efetivos de acordo com as causas de morte.	66
Figura 4.15. Proporção de óbitos por doença do aparelho circulatório nos municípios de Portugal.	67
Figura 4.16. Número de óbitos por acidente de viação nos municípios de Portugal. ...	68
Figura 4.17. Fluxo de dadores face à morada do dador efetivo e respetivo hospital de colheita.	69
Figura 4.18. Número de doentes transplantados por município no ano de 2017.	71
Figura 4.19. Incidência de doentes transplantados face à sua origem nos municípios de Portugal, no ano de 2016.	72
Figura 4.20. Incidência de doentes transplantados face à sua origem nos municípios de Portugal, no ano de 2017.	73
Figura 4.21. Fluxos de acesso de doentes transplantados com rim aos hospitais com unidade de transplantação Renal.	74
Figura 4.22. Fluxos de acesso de doentes transplantados com fígado aos hospitais com Unidade de Transplantação Hepática	75
Figura 4.23. Fluxos do acesso de doentes transplantados com coração aos hospitais com Unidade de Transplantação cardíaca	76
Figura 4.24. Fluxos do acesso de doentes transplantados com pulmão ao hospital com Unidade de Transplantação Pulmonar	76
Figura 4.25. Fluxos do acesso de doentes transplantados com pâncreas aos hospitais com Unidade de Transplantação pancreática	77
Figura 5.1. <i>Clusters</i> e <i>outliers</i> espaciais dos dadores efetivos em Portugal.	80
Figura 5.2. Resíduos da regressão linear múltipla.	83
Figura 5.3. Árvore de regressão para os fatores explicativos dos dadores efetivos.	84
Figura 5.4. <i>Clusters</i> espaciais de doentes transplantados de rim nos municípios de Portugal	91
Figura 5.5. <i>Clusters</i> espaciais de doentes transplantados de fígado nos municípios de Portugal	92
Figura 5.6. <i>Clusters</i> espaciais de doentes transplantados de pulmão nos municípios de Portugal.	93
Figura 5.7. <i>Clusters</i> espaciais de doentes transplantados de coração nos municípios de Portugal.	94

Figura 5.8. <i>Clusters</i> espaciais de doentes transplantados de pâncreas nos municípios de Portugal.	95
Figura 5.9. Resíduos da regressão linear múltipla.....	97
Figura 5.10. Árvore de regressão para os fatores explicativos dos doentes transplantados	98

Índice de quadros

Quadro 3.1. Esquema da aplicação dos métodos de análise espacial	34
Quadro 4.1. Relação dos dadores efetivos por camas de UCI, nos hospitais com neurocirurgia.	57
Quadro 4.2. Relação dos dadores efetivos por camas de UCI, nos hospitais sem neurocirurgia	58
Quadro 4.3. Número de doentes transplantados por hospital e respetivo órgão no ano de 2016.	70
Quadro 4.4. Número de doentes transplantados por hospital e respetivo órgão no ano de 2017.	71
Quadro 5.1. Variáveis explicativas pré-selecionadas	81
Quadro 5.2. Valores de VIF para as variáveis independentes.	82
Quadro 5.3. Parâmetros (estrutura simplificada) de divisão da árvore de regressão dos dadores efetivos	85
Quadro 5.4. Variáveis preditoras (independentes) a utilizar no modelo e respetiva importância.....	85
Quadro 5.5. Análise da <i>performance</i> das cinco melhores Redes Neurais Artificiais para os diferentes conjuntos de variáveis.	87
Quadro 5.6. Análise sensitiva das variáveis.....	88
Quadro 5.7. Índice Global de Moran referente às várias tipologias de doentes transplantados em Portugal.	90
Quadro 5.8. Variáveis explicativas pré-selecionadas.	96
Quadro 5.9. Valores de VIF das variáveis independentes.	96
Quadro 5.10. Parâmetros (estrutura simplificada) de divisão da árvore de regressão dos doentes transplantados.....	99
Quadro 5.11. Variáveis preditoras (independentes) a utilizar no modelo e respetiva importância.....	99
Quadro 5.12. Análise da <i>performance</i> das cinco melhores RNA para os diferentes conjuntos de variáveis.	101
Quadro 5.13. Análise sensitiva das variáveis.....	102

Siglas e Acrónimos

ACSS –Administração Central do Sistema de Saúde

CNT – Coordenação Nacional da Transplantação

DBD – *Donation after Brain Death*

DCD – *Donation after Circulatory Death*

DGT – Direção Geral do Território

EPE – Entidade Pública Empresarial

IRODat – *International Registry on Donation and Transplantation*

INE – Instituto Nacional de Estatística

IPST – Instituto Português do Sangue e da Transplantação

LISA – *Local Indicators of Spatial Association*

RENDA – Registo Nacional de Não Dador

RLM – Regressão Linear Múltipla

RGP – Regressão Geográfica Ponderada

SIG – Sistemas de Informação Geográfica

SIOPT - Sistema de Informação da Organização Portuguesa de Transplantes

SNS – Serviço Nacional de Saúde

TCE – Traumatismo Crânio Encefálico

UCI – Unidade de Cuidados Intensivos

UT – Unidade de Transplantação

VIF - *Variance Inflation Factor*

Agradecimentos

A realização deste relatório de estágio não seria possível sem o contributo de terceiros, pessoas com quem houve uma frequente troca de ideias e um trabalho lado a lado tendo em vista a concretização e o progresso. Assim sendo, deixo aqui o meu agradecimento.

Ao meu orientador, Professor Doutor Jorge Rocha a quem expresso o meu agradecimento por toda a orientação, simplicidade, permanente disponibilidade e sobretudo por todos os ensinamentos.

Ao Professor Dr. Miguel Marques pela simpatia, e possibilidade de frequentar as suas aulas de geomarketing.

À instituição onde realizei o estágio curricular, o Instituto Português do Sangue e da Transplantação, em particular à equipa da Coordenação Nacional da Transplantação. À minha orientadora de estágio, Dra. Catarina Bolotinha por todo o auxílio, disponibilidade para esclarecer dúvidas, correções e boa disposição que sempre demonstrou. À Dra. Ana França, Dr. Paulo Severino, Enf.^a Vanda Palmeiro, Bruna Menezes e Rita Costa agradeço também pelo auxílio, à-vontade e amabilidade com que me receberam ao longo deste estágio.

Agradeço aos meus amigos. Aqueles que, ao longo deste mestrado, tive a fortuna de conhecer e com quem foram partilhados bons momentos. Agradeço particularmente à Letícia F., Susana M., Pedro B., Tiago A., Daniela H. e a todos aqueles que de uma forma mais direta ou indireta sempre estiveram presentes.

Por fim, e não menos importante, aos meus queridos pais por todo o amor, amparo e apoio incondicional em todas as minhas decisões.

A todos vocês, o meu muito obrigada!

Capítulo 1. Introdução

No âmbito do Mestrado em Sistemas de Informação Geográfica e Modelação Territorial aplicados ao Ordenamento realizou-se um estágio curricular com duração de 4 meses no Instituto Português do Sangue e da Transplantação (IPST), mais especificamente no departamento da Coordenação Nacional da Transplantação (CNT). A CNT é responsável pela regulamentação e normalização das atividades de dádiva, colheita e transplantação de órgãos, tecidos e células de origem humana.

Portugal ocupa hoje um lugar de destaque em relação à doação de órgãos mundial. A transplantação é um tratamento eficaz a salvar vidas, por vezes a única esperança para quem aguarda um transplante. Sustentado quer a nível da ciência médica, quer a nível da economia da saúde, contribui para a melhoria das condições de vida da sociedade. Contudo, têm-se disponibilizado poucas ferramentas que possibilitem o estudo geográfico destes fenómenos, em que o apoio das tecnologias de informação geográfica se pode revelar útil.

Este estágio teve por objetivo principal a reflexão sobre os seguintes pontos de atuação:

- i) deteção de potencial de doação nos municípios de Portugal;
- ii) identificação de municípios onde, face às patologias associadas aos doentes transplantados, há a possibilidade de outros afetados pela mesma patologia virem a contribuir para o número de transplantados;
- iii) identificação de fluxos de acesso¹ de dadores de órgãos;
- e iv) identificação de fluxos de acesso de doentes transplantados;

Este trabalho pretende também demonstrar que a localização e a componente territorial desempenham um papel relevante na perceção e explicação de fenómenos ao nível da morbilidade e mortalidade em saúde. É neste sentido que os Sistemas de Informação Geográfica demonstram ser fortes ferramentas, através do armazenamento

¹ Entenda-se a relação entre a distância da residência do dador/recetor ao hospital e o número de dadores/transplantados.

de informação, análise e explicação de fenómenos espaciais, o que por sua vez dá ênfase à potencialidade dos mapas.

O presente relatório encontra-se dividido em duas componentes. Numa primeira instância, realizou-se um levantamento e tratamento não só da informação disponibilizada no *website* do Instituto Nacional de Estatística (INE), como também da que foi cedida pelo IPST ao longo do estágio, de modo a que fosse possível a criação de uma base de dados. Posteriormente, foi realizada a inventariação e georreferenciação dos hospitais do Serviço Nacional de Saúde (SNS), das unidades de transplantação, assim como a geocodificação da morada dos dadores e dos doentes transplantados.

Preparada a informação, procedeu-se à realização da segunda componente do estudo, através de análises ligadas à estatística espacial de modo a compreender as dinâmicas espaciais da doação e da transplantação em Portugal. Dada a informação obtida, realizou-se uma análise preliminar e espacial da doação e transplantação, cruzando variáveis de cariz populacional. *A posteriori*, aplicaram-se métodos estatísticos, nomeadamente o Índice Global e Local de Moran e análise de Hot Spot, com vista a identificar padrões espaciais e constatar se a distribuição espacial do número de dadores e doentes transplantados é aleatória ou aglomerada (em *clusters*). Foi também realizada uma análise de potencial, recorrendo ao método Regressão Linear Múltipla, como forma de perspetivar se alguns aspetos populacionais influenciam o número de dadores, ou se existe alguma especificidade na associação de patologias a quem necessita de transplante, e se sim, em que locais. Por fim, recorreu-se às Redes Neurais com o intuito de confrontar modelos, assim como perspetivar face às variáveis predefinidas, qual a rede e quais as variáveis que melhor justificam a localização dos doentes transplantados.

1.1. Doação e Transplantação em contexto Mundial

A transplantação de órgãos demarca-se como uma das mais notáveis conquistas científicas, assim como um dos maiores progressos do século passado, no que respeita a cuidados de saúde. Concebida através de uma série de estudos experimentais, tornou-

se numa realidade viável apenas no século XX, altura em que diversos foram os marcos históricos que transverteram a transplantação no que é hoje, uma terapêutica de relevância dos cuidados de saúde.

O termo “transplante” foi utilizado pela primeira vez pelo cirurgião John Hunter, em 1771. Destaquem-se os marcos históricos da transplantação ao longo do século XX: o primeiro transplante de rim que teve lugar em 1954, em Boston, quando os cirurgiões Joseph Murray e David Hume extraíram um rim de um gêmeo verdadeiro, para implementá-lo no corpo do seu irmão, tendo este último sobrevivido oito anos após a cirurgia; o primeiro transplante hepático levado a cabo em 1963, por Starzl, também nos EUA; o primeiro transplante de pulmão por James Hardy, neste mesmo ano e país; o primeiro transplante de pâncreas em 1966, executado por Richard Lillehei e Kelley em Minneapolis; e, por fim, o primeiro transplante de coração, que data de 1967, operado por Chistian Barnard na Cidade do Cabo, na República Sul-Africana (Costa, 2017).

Na atualidade é possível definir a transplantação de órgãos como um procedimento no qual se implementa um órgão ou tecido proveniente de um dador para um recetor, cuja finalidade é a melhoria das condições de vida, o que na generalidade dos casos consiste na única esperança (Miedes, 2008). O transplante de órgãos é a melhor e, muitas vezes, a única alternativa terapêutica para um grande número de doentes com falência terminal de um órgão. Porém, tal procedimento só é possível devido ao desenvolvimento paralelo de um programa de doação de órgãos.

Segundo o IRODaT², atualmente existem 103 países no mundo que realizam atividade de transplantação, porém, apenas 9,7% destes obtiveram bons resultados de doação. Portugal tem sido, nos últimos anos, um dos países com maiores taxas de doação por milhão de habitantes (pmh), com taxas de dadores falecidos só ultrapassadas por Espanha, o que coloca Portugal na vanguarda da doação a nível mundial.

² *International Registry in Organ Donation and Transplantation, March 2018 (in press)*

Em comparação com os restantes países do mundo, em 2017 (Figura 1.1), Portugal foi o segundo país com maior número de dadores falecidos (com 34,1 dadores pmh), sendo o primeiro lugar ocupado por Espanha, seguindo-se a Bélgica, Croácia, EUA, França, Itália, República Checa, entre outros (IRODaT, 2017).

Figura 1.1. Doação de órgãos em contexto europeu por milhão de habitantes no ano de 2017 (fonte: *Newsletter Transplant*, 2018)

O acesso de doentes à transplantação de órgãos varia de acordo com as situações nacionais, sendo parcialmente determinado pelo custo dos cuidados de saúde, suporte técnico e, mormente, a disponibilidade de órgãos. A escassez de órgãos é um problema universal. Em alguns países, principalmente nos países em desenvolvimento, a implementação de um programa de doação em dador falecido é dificultada por fatores

socioculturais, legais, entre outros. Relativamente a países ditos desenvolvidos, onde as taxas de doação de órgãos tendem a ser maiores, a colheita de órgãos não consegue dar resposta à crescente demanda (Shimazono, 2007). O recurso à doação por dador vivo para transplante renal e hepático também é exercido um pouco por todo o mundo. Devido à grande escassez de órgãos, a comercialização destes ou o tráfico de pessoas com o objetivo de obter órgãos para transplante tem-se revelado uma preocupação acrescida.

1.2. Transplantação de Órgãos em Portugal

Em Portugal a transplantação conta já com quatro décadas de atividade contínua. Em perspetiva histórica, destaque-se o singular e relevante episódio que ocorreu a 20 de julho de 1969. No mesmo dia em que curiosamente um ser humano chegava pela primeira vez à lua, o Professor Doutor Linhares Furtado³, em Coimbra, antecipando-se ao futuro da evolução técnico-científica neste domínio, realizou com sucesso a primeira transplantação de rim proveniente de um dador vivo aparentado (Costa, 2017).

Tal ato não se veio a reiterar durante vários anos, apesar da sua prática já ser frequente noutros países. O contexto nacional à época não era o mais promissor para a transplantação, entre muitas outras razões, destacavam-se a inexistência de centros de histocompatibilidade, de legislação que regulasse a colheita e o transplante de órgãos, assim como, uma rede de hemodiálise incipiente (Oliveira, 2011). Só em 1980 é que se veio a fazer a primeira colheita de órgãos vitais, abrindo caminho a uma atividade contínua e sistematizada, de maneira a possibilitar a perpetuação da vida humana que em essência é a tônica da ciência médica.

Passados mais de 30 anos, a transplantação em Portugal deteve um trajeto notável, através da realização dos mais diversos transplantes. Ao rim sucederam-se outros órgãos, nomeadamente o primeiro transplante cardíaco em Lisboa, em 1986, realizado

³ Professor Catedrático Jubilado da Faculdade de Medicina da Universidade de Coimbra. Considerado o pioneiro da transplantação de órgãos sólidos em Portugal.

pelo cirurgião Queiroz e Melo, ocorrendo as primeiras transplantações hepáticas em 1992, no ano de 1993 dá-se a primeira transplantação pancreática e, em 1987, tem lugar a primeira transplantação pulmonar.

1.3. Doação de Órgãos de dador falecido

A doação de órgãos é um ato generoso e altruísta em que uma pessoa manifesta voluntariamente em vida o seu arbítrio de doar, após a morte, uma ou várias partes do seu corpo para transplante (Rubio e Alba, 2000).

Atualmente e como resultado dos últimos anos, Portugal tem ocupado os lugares cimeiros na doação de órgãos, tanto a nível europeu como mundial. Porém, todos os anos ainda morrem pessoas em lista de espera para transplante.

Segundo o despacho nº 5480/2017, de 23 de junho de 2017, *“a falta de identificação e referenciação de possíveis dadores é a principal causa de baixos níveis de doação designadamente em hospitais com elevado potencial de doação, devendo ser precoce e obrigatoriamente identificadas todas as oportunidades, promovendo-se a avaliação caso a caso”*.

Uma atividade de doação, estável e consolidada é fundamental para garantir uma resposta às necessidades dos doentes que aguardam por um transplante (IPST, 2017⁴). A carência de órgãos para transplantação, que existe à escala global, deve ser contrariada de forma dinâmica, com a implementação de medidas corretivas e adequadas em cada momento, tendo em conta a evolução das circunstâncias epidemiológicas, demográficas e sociais (Despacho nº 5480/2017, de 23 de junho de 2017).

A deteção de potenciais dadores é o ponto de partida para a transplantação. À luz da lei em vigência, em Portugal são considerados como potenciais dadores *post mortem*

⁴ Site: www.ipst.pt –/ Acedido a 18/04/2018

“todos os cidadãos nacionais, apátridas e estrangeiros residentes em Portugal que não tenham manifestado, junto do Ministério da Saúde, a sua qualidade de não dadores” (Artigo 10º da Portaria nº 22/2007, de 29 de junho).

A total ou parcial indisponibilidade para a dádiva *post mortem* de certos órgãos ou tecidos para determinados fins é manifestada junto do Ministério da Saúde, através da inscrição no RENNDA⁵.

Atualmente, o cadáver humano é a principal fonte de órgãos. O possível dador de órgãos traduz-se num *“doente com uma lesão ou doença cerebral catastrófica, ou um doente com insuficiência cardíaca e, aparentemente, medicamente apropriado para a doação de órgãos”* (Direção Europeia da Qualidade dos Medicamentos e Cuidados de Saúde do Conselho da Europa, 2016), e no que concerne a dadores falecidos de órgãos, estes dividem-se em duas categorias:

- Doação após morte cerebral (DMC) (DBD do inglês *Donation after Brain Death*);
- Doação após morte cardiocirculatória (DPC) (DCD do inglês *Donation after Circulatory Death*).

A morte cerebral, conhecida também por morte neurológica, traduz-se na impossibilidade de consciência, ou seja, é uma morte irreversível. A doação após morte cerebral, segundo o manual de coordenação de transplantes (2014), ocorre em pacientes cuja morte é diagnosticada após sofrerem uma lesão cerebral grave de acordo com critérios neurológicos, enquanto as funções respiratórias e cardiovasculares são mantidas artificialmente. Tais características levam a potenciais dadores de órgãos, tais como: rins, fígado, coração, pulmões e pâncreas.

⁵ “Tem por finalidade organizar e manter atualizada (...) a informação relativa à indisponibilidade para a colheita de órgãos ou tecidos. Os dados pessoais destinam-se à verificação, antes de iniciada a colheita, por parte das entidades que nos termos da lei aplicável procedem à colheita *post mortem* de órgãos ou tecidos, da existência de oposição”. site: www.spms.min-saude.pt/product/rennda/

Relativamente à doação após a morte cardiocirculatória, esta sucede em pessoas que apresentam uma completa e irreversível cessação de toda a função circulatória e respiratória, com a consequente certificação de morte⁶.

Segundo a legislação, despacho nº 14341/2013, 29 de outubro de 2013, os dadores após paragem cardiocirculatória são categorizados segundo critérios de Maastricht, em cinco categorias:

- Categoria I – Morte à chegada – representa as vítimas de acidente ou de paragem cardiocirculatória, encontradas mortas no local pela equipa de reanimação (...);
- Categoria II - Reanimação Infrutífera - doentes que tiveram morte cardíaca súbita ou com lesões cerebrais irreversíveis seguidas de paragem cardiocirculatória, ocorridas no hospital ou no exterior deste, a quem se realizaram manobras de reanimação sem êxito;
- Categoria III - Doentes irreversíveis que aguardam paragem cardíaca - doentes que padecem de uma patologia incurável e de evolução irreversível, internados em Serviços de Medicina Intensiva, aos quais se suspende o suporte das funções vitais;
- Categoria IV - Paragem após diagnóstico de morte cerebral;
- Categoria V - Morte por paragem cardíaca em doentes hospitalizados.

Citando a alínea 3 do despacho nº 5480/2017, de 23 de junho de 2017 *“através das normas hospitalares de doação o grupo de trabalho deve ter em conta os circuitos dos possíveis dadores, focalizando-se no serviço de urgência, serviços de medicina intensiva, unidades de acidentes vasculares cerebrais e de neurocríticos, assim como outras unidades de internamento”*.

⁶ Disponível no “Guia para a qualidade e segurança dos órgãos para transplantação, 2013”

1.4. Instituto Português do Sangue e da Transplantação, IP

O IPST é um organismo público com jurisdição sobre todo o território nacional na área da medicina transfusional e da transplantação.

É missão do IPST, IP *“garantir e regular, a nível nacional a atividade da medicina transfusional e da transplantação, e garantir a dádiva, colheita, análise, processamento, preservação, armazenamento e distribuição de sangue humano, de componentes sanguíneos, de órgãos, tecidos e células de origem humana”* (IPST, 2018). A visão do instituto é a promoção da doação, com o objetivo de possibilitar a melhoria da sobrevivência e qualidade de vida dos seres humanos, tendo por base as evidências das melhores práticas e conhecimentos científicos.

O Plano Estratégico de 2014-2016⁷ do IPST refere um conjunto de medidas plurianuais de implementação, sendo uma delas a “utilização das Tecnologias de Informação e Comunicação de modo a potenciar mudança e modernização administrativa através da integração dos sistemas de informação, aumentando a sua fiabilidade e interoperabilidade”, onde este projeto se enquadra.

Sendo assim, a carência de tratamento de informação, bem como de uma base de dados digital ou de informação georreferenciada relativamente à doação e transplantação a nível nacional pode causar alguns constrangimentos. O presente estágio assentou, por isso, na criação dessa base de dados, por sua vez, objeto de um tratamento de informação.

Segundo o último Plano Estratégico de 2015-2024, as orientações estratégicas na área funcional da transplantação passam por:

- *Aumentar o número de órgãos e tecidos disponíveis para transplantação;*

⁷ Site: www.ipst.pt – Acedido a 04/05/2018

- *Reorganizar a Rede de Coordenação da Colheita e de Transplantação, com definição do número de GCCT e de Unidades de Transplantação, bem como a sua articulação.*
- *Implementar o Registo Português da Transplantação a nível nacional, com integração das diferentes bases de dados na área da transplantação;*
- *Reformular o modelo de relacionamento com as associações e grupos de dadores.*

1.5. A Coordenação Nacional da Transplantação

A Coordenação Nacional da Transplantação (CNT) é uma *“unidade orgânica do Instituto Português do Sangue e da Transplantação, responsável pela regulamentação e normalização das atividades de dádiva, colheita e transplantação de órgãos, tecidos e células de origem humana”* (IPST, 2018). A sua estrutura baseia-se numa organização a três níveis:

- Nacional (Coordenação Nacional da Transplantação - CNT);
- Regional (Gabinetes Coordenadores de Colheita e Transplantação - GCCT);
- Local (Coordenadores Hospitalares de Doação, Unidades de Colheita de Tecidos e Células, Unidades de Transplantação/aplicação de órgãos tecidos e células).

O objetivo principal da CNT é a promoção da doação altruísta, promovendo o aumento progressivo da disponibilidade de órgãos, tecidos e células para transplante, garantindo, desta forma, que todos os cidadãos que necessitem de um transplante tenham as maiores e melhores possibilidades de o conseguir. Para tal, a CNT integra uma equipa de profissionais, em diferentes departamentos de modo a que tal pressuposto seja possível (Figura 1.2).



Figura 1.2. Estrutura orgânica da Coordenação Nacional da Transplantação.

A rede nacional de coordenação de colheita e transplantação, adiante designada por rede, é constituída pelos Coordenadores Hospitalares de Doação e pelos GCCT.

As Normas Hospitalares de Doação constituem uma obrigação de todas as instituições de saúde da rede hospitalar, tanto do Serviço Nacional de Saúde (SNS) como da Rede de Hospitais Privados devidamente autorizados no âmbito do artigo 1º da Portaria nº 76/2014, de 21 de março e de acordo com a Lei nº 36/2013, de 12 de junho. Porém, no presente estudo apenas serão abordados os hospitais do SNS.

A atividade de colheita de órgãos, em Portugal, encontra-se repartida em termos geográficos por cinco GCCT, localizados nos seguintes hospitais:

- a) Hospital de São José, Centro Hospitalar de Lisboa Ocidental, EPE;
- b) Hospital de Santa Maria, Centro Hospitalar de Lisboa Norte, EPE;
- c) Hospital de Santo António, Centro Hospitalar do Porto, EPE;
- d) Hospital de São João, EPE;
- e) Hospitais da Universidade de Coimbra.

Relativamente às Unidades de Transplantação ao nível Local, estas operam ao nível da transplantação renal, hepática, pancreática, cardíaca e pulmonar distribuindo-se pelos diversos centros hospitalares. No que respeita à Unidade de Transplantação **Renal**, esta opera nos seguintes hospitais:

- Centro Hospitalar do Porto;
- Centro Hospitalar de São João;
- Centro Hospitalar de Coimbra;
- Centro Hospitalar de Lisboa Central;
- Hospital Garcia de Orta;
- Centro Hospitalar de Lisboa Norte;
- Centro Hospitalar de Lisboa Ocidental.

Unidade de Transplantação **Hepática:**

- Centro Hospitalar do Porto;
- Centro Hospitalar Universitário de Coimbra;
- Centro Hospitalar de Lisboa Central.

Unidade de Transplantação **Pancreática:**

- Centro Hospitalar do Porto;
- Centro Hospitalar de Lisboa Central.

Unidade de Transplantação **Cardíaca:**

- Centro Hospitalar de São João;
- Centro Hospitalar Universitário de Coimbra;
- Centro Hospitalar de Lisboa Central;
- Centro Hospitalar de Lisboa Ocidental.

Unidade de Transplantação **Pulmonar:**

- Centro Hospitalar de Lisboa Central.

1.6. Área de Estudo

A área onde se desenvolveu o caso de estudo corresponde a Portugal Continental e às regiões autónomas da Madeira e dos Açores (Figura 1.3). Portugal apresenta uma área de aproximadamente 92 090 km², e, de acordo com os censos de 2011, um total de 10 562 178 habitantes. Relativamente a Portugal Continental, este dispõe de uma área total de 89 281 km², 18 distritos e 278 concelhos. Quanto às duas regiões autónomas insulares (Madeira e Açores), a Madeira detém 11 municípios e os Açores 19 municípios.

Em Portugal subsistem ao nível regional cinco GCCT, que se articulam através dos Coordenadores Hospitalares de Doação nos 48 hospitais do SNS. No que respeita às unidades de transplantação, estas situam-se em 8 hospitais de Portugal Continental.

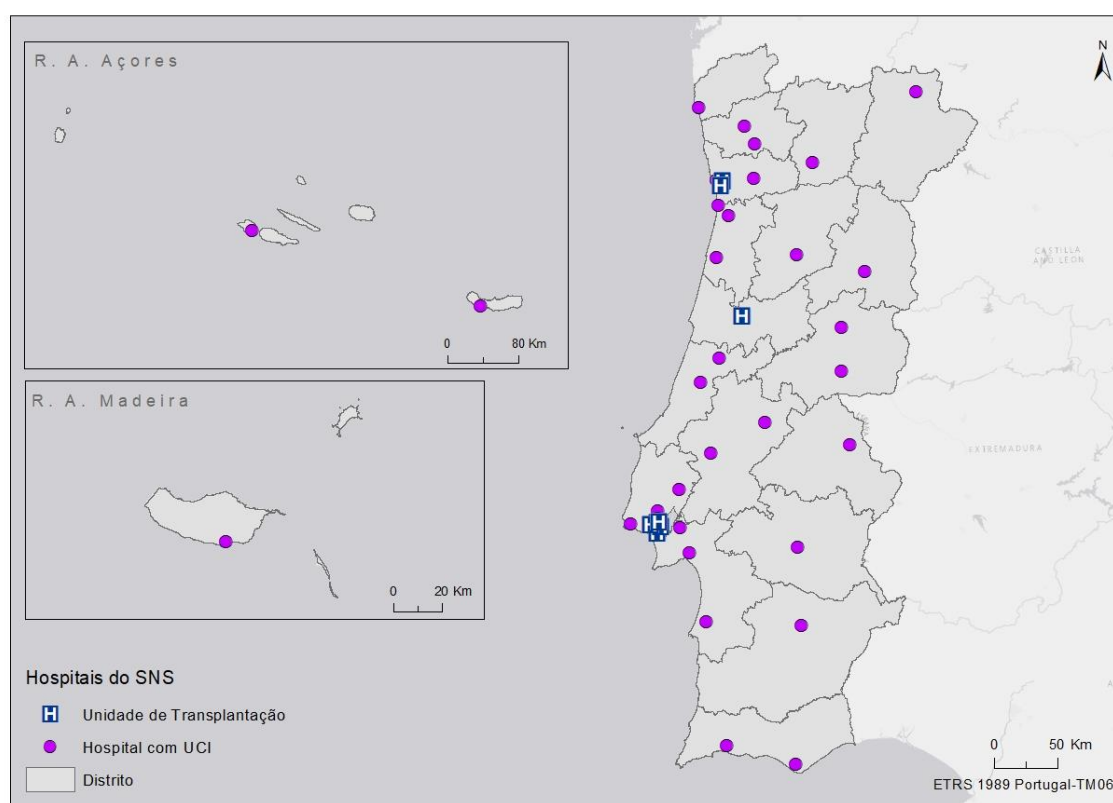


Figura 1.3. Unidades de transplantação e hospitais com Unidade de Cuidados Intensivos (UCI).

Capítulo 2. Geografia e SIG

A cartografia, como ciência e arte de construir mapas, opera como intérprete, traduzindo fenómenos do mundo real (dados primários) em representações corretas, claras e compreensíveis para nosso uso (Burrough e McDonnell, 1998).

No final dos anos 1960, era introduzida pela primeira vez nas Universidades Americanas a cartografia analítica, através de Waldo Tobler. Tal acontecimento foi motivado pela sua visão de que os geógrafos utilizavam mapas como ferramentas analíticas para compreender e teorizar acerca da Terra e da distribuição de fenómenos à sua superfície (Kemp, 2008).

Também nesta década surgem os SIG, embora na época revelassem problemas técnicos significativos, e.g. converter o conteúdo de mapas em formato analógico para digital. Não obstante, com o passar dos anos, tais entraves foram-se aprimorando, os SIG foram-se expandindo e problemas de computação foram resolvidos. Por volta de 1995, a visão complexa dos SIG foi comutada, constituindo agora o meio através do qual o Homem comunica conhecimentos sobre a superfície da Terra (Shekar e Xiong, 2008).

Hoje, são diversas as áreas do conhecimento que consideram a variável espaço um elemento explicativo. Tais áreas têm vindo a beneficiar de um vasto conjunto de técnicas de geoprocessamento, através do tratamento e manipulação de dados geográficos, que vão desde a deteção remota, aos sistemas de posicionamento global (GPS), cartografia analítica, inclusive a mais complexa técnica de geoprocessamento – os SIG (Nossa, 2005).

Há vários anos que as ações de saúde pública se baseiam na análise e uso de dados espaciais, onde a geografia e os SIG podem revelar-se uma mais-valia nos estudos em saúde (Hanchette, 2003).

2.1. Geografia da Saúde

A intervenção da Geografia da Saúde permite ler, interpretar factos, e inclusive, estabelecer hipóteses explicativas e de ação sobre alguns dos problemas existentes, tendo por base uma matriz territorial (Santana, 2014).

“Quem se aventura a fazer Geografia da Saúde tem de compreender as necessidades dos técnicos de saúde, da população, dos gestores e sobretudo oferecer alternativas de análise e de intervenção” (Barcellos, 2008).

A sua contribuição em contexto de equipas multidisciplinares pode mostrar-se profícua. Esta área ocupa uma posição nodal, como sendo um espaço onde convergem ou se cruzam fenómenos naturais, socioeconómicos, culturais e comportamentais, de importância primordial na explicação de padrões de saúde e doença (Santana, 2014).

O interesse no estudo das relações entre espaço, lugar e saúde recua a épocas remotas. A primeira tentativa de construção de um mapa de doenças realizou-se nos Estados Unidos da América por médicos que cartografaram a residência de pessoas afetadas com o vírus da febre-amarela (Seaman, 1798). Também no início do século XVIII surgiam os *dot maps* (mapas de pontos que mostravam o padrão da epidemia), sendo os mais célebres os de Seaman (mapas de febre-amarela num setor de Nova Iorque, em 1798) (Nogueira e Remoaldo, 2010).

Segundo Nogueira (2007), a Geografia da Saúde pode ser definida como uma *“subdisciplina geográfica, em que a sua especificidade advém, por um lado, da incorporação da dimensão espacial no estudo da saúde, (abrangendo campos que vão desde os resultados em saúde ao planeamento dos serviços de saúde) e, por outro lado, da aplicação de ferramentas e abordagens da Geografia ao seu objeto de estudo”*.

Em síntese, o objetivo geral da Geografia da Saúde é proporcionar conhecimentos por forma a entender as relações que se estabelecem entre (1) as condicionantes da saúde,

(2) as políticas e a organização dos serviços na saúde das populações e (3) as consequências de todos estes no desenvolvimento territorial (Santana, 2014).

2.2. Sistemas de Informação Geográfica

Os SIG datam dos anos 60, quando foram realizadas as primordiais tentativas de aplicação de sistemas informáticos ao tratamento e processamento de informação geográfica. Os SIG podem ser definidos de distintas formas, e inúmeras são as possibilidades e ferramentas de tratamento e análise de informação disponibilizada por estes. É possível afirmar que não existe uma definição universal para os SIG (Taylor e Overton, 1991), mas antes várias definições, umas mais elementares, outras mais complexas.

Kemp, *et al.* (2008) definem os SIG como um sistema de incorporação de *software*, capaz de executar um conjunto de funções, onde a informação é efetuada por uma comunidade de pessoas de áreas científicas diversas. Tal informação é representada em formato digital para que possa ser trabalhada, visando a utilização dos SIG na resolução de problemas, numa ótica de avanço científico. Segundo Ozernoy (1981), os SIG representam um conjunto de funções automatizadas, que fornecem aos profissionais de saúde capacidades avançadas de armazenamento, acesso, manipulação e visualização de informação georreferenciada. Contudo, Arbeit (1992) define os SIG como *“um sistema computacional que armazena e relaciona atributos não gráficos ou dados geograficamente referenciados com características gráficas, de modo a permitir uma vasta gama de processamento de informação e operações de exibição, bem como a produção de mapas, análise e modelagem”*.

As diferentes definições têm enfoque em dois aspetos: tecnologia e resolução de problemas. A abordagem tecnológica define o SIG como um conjunto de ferramentas para entrada, armazenamento, gestão, análise, e finalmente, saída de dados com forte componente espacial (Figura 2.1). A abordagem à resolução de problemas realça o uso dos SIG no apoio à decisão em questões espaciais (Malczewsky *et al.*, 1997).



Figura 2.1. Componentes de um SIG.

2.3. Relevância dos SIG na Saúde Pública

A associação entre a medicina e a geografia remota a épocas antigas. Muito antes da era da informação e do desenvolvimento, à data de 1948, John Snow⁸ demonstrou, na cidade de Londres, a importância da geografia nas atividades de monitorização e análise.

Na época, Londres vivia assolada por um violento surto, marcado pela ocorrência de 500 mortes em apenas 10 dias. Dada tal realidade, Snow deliberou mapear as mortes por cólera⁹ (bactéria *vibrio cholerae*) em cerca de seis quarteirões da cidade de Londres (Snow, 1855). Aquando do mapeamento, facilmente sobressaíram *clusters*¹⁰ espaciais com efeitos de decadência face à distância das fontes de água (Moore e Carpenter, 1999). Tal mapa (Figura 2.2) conduziu à hipótese que o "campo da cólera", como Snow designava, possuía como epicentro uma bomba de água que abastecia cerca de seis quarteirões da cidade de Londres, mais especificamente a *Broad Street*. Após a remoção da bomba, observou-se uma redução significativa na incidência de cólera e na sua mortalidade (Cerdeira e Valdivia, 2007), o que auxiliou a comunidade médica a concluir que a cólera era uma doença relacionada com a água, ao invés de surgir por miasma¹¹ como se acreditava generalizadamente na época (Lumpkin, 2003).

Mesmo sem conhecer a causa bacteriana ou o meio de transmissão da cólera em meados do século XIX, Snow conseguiu conter a propagação do surto, uma vez que

⁸ Médico inglês considerado o pai da epidemiologia moderna.

⁹ Infecção intestinal causada por uma bactéria e transmitida pela ingestão de água e comida contaminada.

¹⁰ Um *cluster* espacial é definido como "o excesso de eventos ou de valores num local determinado que difere significativamente da variação geográfica esperada" (Jacquez, 2008).

¹¹ Emissão proveniente de detritos orgânicos em decomposição, considerada outrora (antes dos avanços da microbiologia) como causadora de doenças e epidemias. (Dicionário de Língua Portuguesa. Porto Editora, 2003-2018).

compreendeu os seus aspetos espaciais (Moore e Carpenter, 1999). Tal caso remete-nos à primeira lei da geografia de Tobler (1970), em que “tudo está relacionado com tudo, mas os fenómenos mais próximos estão mais relacionados do que os fenómenos distantes”.



Figura 2.2. Casos de cólera durante a epidemia em Londres data de 1854 (Gilbert, 1958).

Esta é uma situação típica, dado que a relação espacial entre os dados contribui significativamente para o avanço na compreensão do fenómeno, revelando-se um dos primeiros exemplos de análise espacial. Também Bickmore e Tulloch (1979) destacam o

enorme potencial da cartografia para ajudar a compreender de que modo emergem e/ou se agravam certas patologias.

Os sistemas computacionais, nomeadamente a produção massiva de mapas e gráficos de compreensão simples, têm-se revelado uma ferramenta útil na medicina, sobretudo na epidemiologia e no planeamento da prestação dos cuidados de saúde. Entender a distribuição espacial e temporal da doença é parte integrante da investigação em saúde pública (Rushton 2003; Ricketts, 2003). Os métodos de análise quantitativa e estatística através de um SIG facilitam a análise de uma ampla gama de determinantes, incluindo fatores demográficos, socioeconómicos, geográficos e ambientais que influenciam a transmissão de doenças (Blossom , *et al.*, 2011).

Recentemente o recurso aos SIG e a representação espacial de diversas questões relacionadas com a saúde auxiliam os profissionais na tomada de decisões de modo eficiente e célere, em particular no campo da saúde pública (Fradelos, *et al.*, 2014).

Segundo Elzubair (2009), as principais vantagens do uso dos SIG no domínio da saúde pública são:

- A aptidão de combinar diferentes conjuntos de dados provenientes de diversas fontes e geri-los de forma adequada, permitindo um melhor planeamento das intervenções no domínio da prevenção da doença e promoção da saúde;
- A diversidade de novas ferramentas de tratamento e análise de dados de índole diversa, tais como, sistemas de posicionamento global, imagens de satélite ou fotografia aérea, revelando-se bastante pertinentes na identificação e explicação da origem de certas patologias;
- A facilidade com que permitem a partilha de dados e com que importam informação proveniente de outros *softwares*, o que no sistema de saúde é deveras relevante;
- O eficiente apoio na tomada de decisão, estimulando a comunicação;
- A capacidade de assegurar a manutenção dos registos guardados, assim como a gestão da informação geográfica, restringindo o acesso da forma desejada.

Atualmente há já diversas organizações internacionais que lidam com a saúde e que procuram conhecer e relacionar aspetos físicos e socioeconómicos com as patologias. Em Portugal o caso mais evidente é a plataforma SIG do Alto Comissariado da Saúde¹², que ilustra, inequivocamente, a relevância e o contributo dos SIG neste domínio, sendo os fatores de índole geográfica determinantes tanto para compreender as consequências, como para antecipar e prevenir impactes negativos, isto é, na delineação de estratégias mais eficientes (Goodchild, 2003).

Uma síntese visual de grandes quantidades de informação abstrata numa imagem alusiva a uma área geográfica é de particular relevância para os profissionais do IPST. Tal visualização possibilita a perceção da distribuição geográfica, no caso do presente estudo, dos dadores de órgãos e dos doentes transplantados em Portugal, possibilitando a criação de vínculos na associação a patologias, entre outros.

Nestes casos, o SIG surge frequentemente como uma ferramenta útil e eficiente na análise de padrões e processos espaciais, que auxilia na compreensão da teia de relações envolvidas (Fonseca, 2012). Contudo, têm-se disponibilizado poucas ferramentas que possibilitem o estudo geográfico destes fenómenos. Segundo a Organização Mundial de Saúde (2010) “o SIG é altamente adequado para análise de dados, revelando tendências e inter-relações que seriam difíceis de encontrar em formato de tabelas”.

A utilidade destas ferramentas na saúde para gestão de informação dependerá em grande parte da qualidade dos dados e dos atributos espaciais. Uma vez que os dados de saúde geralmente advêm de fontes oficiais, podem revelar-se imprecisos e incompletos, o que se traduz na maioria dos casos em dados que chegam com problemas associados, tais como subnotificação, erros de codificação e erros de diagnóstico (Moore e Carpenter, 1999). Atualmente, ainda existe uma grande quantidade de dados de saúde pública gerados manualmente, e, portanto, sujeitos a lapsos. É de relevância a sua informatização, assim como o tratamento e atualização destes.

¹² Disponível na internet em: <http://www.websig.acs.min-saude.pt/>

2.4. Sistema Espacial de Apoio à Decisão

Considerando as noções relacionadas com a teoria da decisão e a generalidade dos problemas de decisão espaciais, a integração de informação geográfica e a utilização de uma aplicação *Web*, fornecida pela *Mapidea location analytics* ao IPST, que assenta no conceito de Sistema Espacial de Apoio à Decisão (SEAD).

“A importância da informação geográfica e da localização é reconhecida no apoio à decisão e tem ganho nos últimos anos, com a democratização do acesso web à informação geográfica, via projetos como o Google Maps ou o Bing Maps uma importância acrescida” (Mileu, 2016).

O SEAD tem origem em duas fontes distintas – os Sistemas de Apoio à Decisão (SAD) e os SIG. Ao longo dos anos, os SAD têm sobressaído enquanto área ativa de pesquisa em SIG, no entanto, a sua capacidade de suportar dados espaciais revelou-se sempre uma das suas principais limitações. Por um lado, os SIG são eficientes na gestão e armazenamento de dados espaciais, mas, por outro, são escassos em ferramentas que facilitem a gestão da tomada decisão. A fusão destas duas tecnologias emergiu no que se entende hoje como SEAD, onde o poder analítico de decisão dos SAD se articula com os recursos espaciais dos SIG (Sugumaran e Sugumaran, 2005).

Um SEAD integra funcionalidades dos SIG, em particular a gestão de dados espaciais, a exibição cartográfica, além do recurso a modelação analítica, uma interface de utilizador flexível e estruturas de dados espaciais complexas (Goodchild, 2003). Pode ser definido como um sistema interativo baseado num computador, que auxilia os decisores na utilização de dados e modelos sendo a finalidade a resolução de problemas não estruturados (Gorry e Morton, 1971).

O SEAD, inicialmente, desenvolveu-se na década de 1970, seguindo-se uma segunda fase entre os anos de 1999 e 2000 (Figura 2.3). Todavia, a sua fase de expansão apenas se veria a revelar após o ano de 2000 (Sugumaran e DeGroote, 2011).

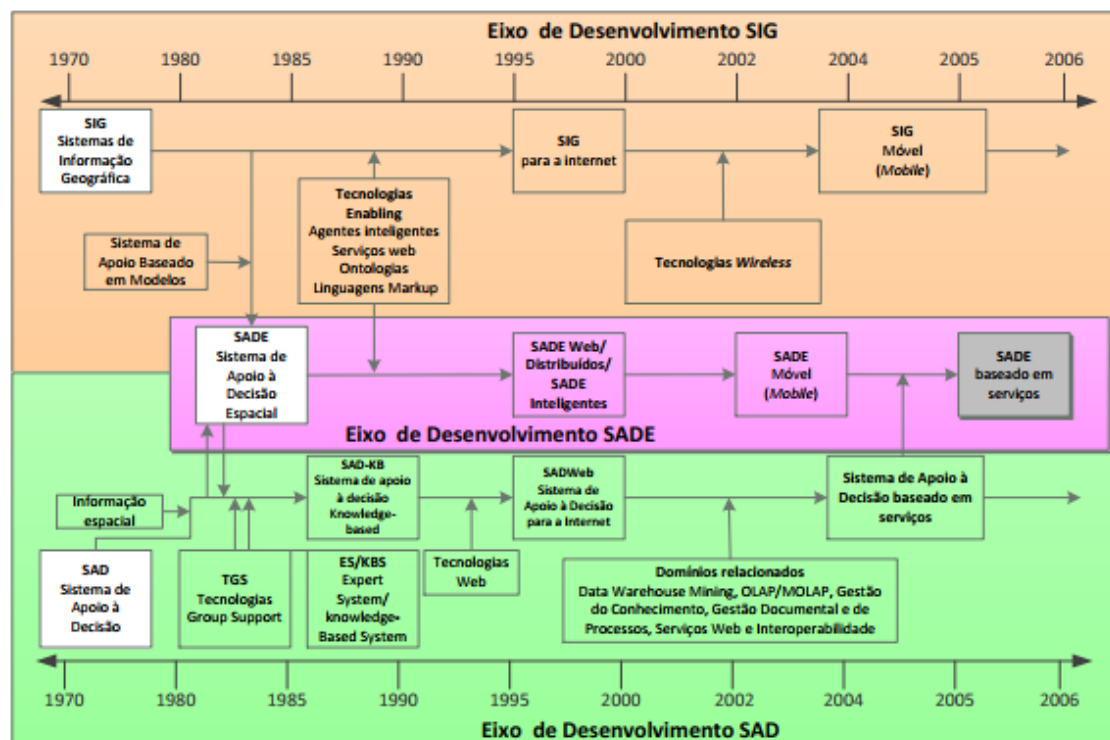


Figura 2.3. Desenvolvimento do SEAD (adaptado de Sugumaran e Sugumaran, 2005:4).

A essência do SEAD inclui um conjunto de características, entre as quais se destacam, segundo Cioca e Cioca (2010):

- A redução de esforços, visando aumentar a capacidade dos decisores, onde o objetivo não é a sua substituição ou transformação em meros agentes que concebem soluções mecanicamente fornecidas pelo computador;
- A finalidade de abordar problemas semiestruturados, onde os decisores utilizam o seu próprio raciocínio no controlo e processo de decisão.

2.4.1. Componentes e tipologias do SEAD

O SEAD deve ser projetado de modo a simplificar o uso, fornecer soluções através de uma panóplia de alternativas flexíveis de fácil adaptação, assim como apoiar métodos analíticos. Para alcançar tais atributos, é fulcral que o SEAD integre certos componentes.

A um nível mais básico subsistem três componentes principais: base de dados, modelo e interface do utilizador.

Malczewski (1997) identifica (1) o Sistema de Gestão de Base de Dados (SGBD), uma vez que contém as funções de gestão da base de dados (geográficas ou não); (2) o Sistema de Gestão do Modelo de Base (SGMB), que contém as funções para gerir o modelo de base; e por último, (3) o Sistema de Gestão e Geração de Diálogos (SGGD), que gere o interface entre o utilizador e o restante sistema (Figura 2.4).

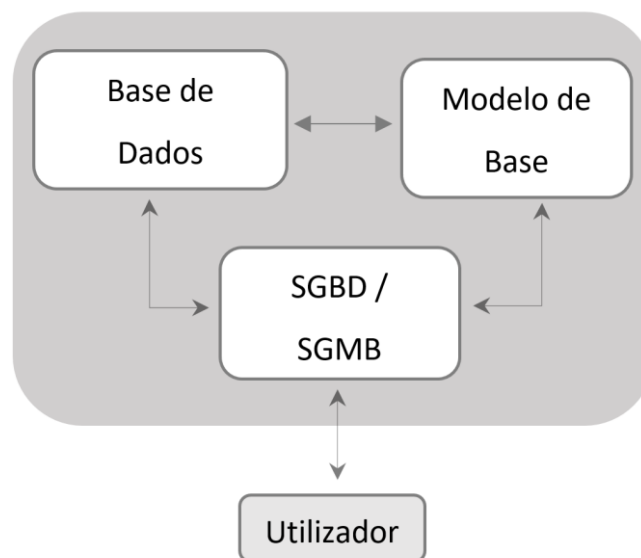


Figura 2.4. Componentes de um SEAD. (adaptado de Maczewski, 1997:4).

Segundo Power (2007), ao nível concetual é possível identificar seis categorias de SAD:

1. SAD orientados a Dados;
2. SAD orientados a Modelos;
3. SAD orientados à Comunicação;
4. SAD orientados a Documentos;
5. SAD orientados à Inteligência Artificial;
6. SAD suportados pelo *Web*.

Com particular relevância para a presente investigação, importa destacar (1) os SAD orientados a dados, ou seja, destinados à manipulação e acesso de grandes quantidades de dados estruturados e séries temporais de fontes internas e externas; e (2) os SAD suportados pela *Web*, sistemas capazes de disponibilizar dados e modelos apropriados aos decisores recorrendo a um navegador de Internet.

2.4.2. WebSEAD

A *Web* revolucionou o desenvolvimento de aplicações. A sua natureza omnipresente e a possibilidade de uma interface simplificada para os utilizadores levaram à implantação de aplicações complexas, como é o caso do SEAD na *Web* (Keenan, 2003).

Os SIG baseados na *Web* combinam o poder da *World Wide Web* com funções de SIG, tais como gerar, visualizar, interagir com mapas, ou executar consultas e análises espaciais. Deste modo, a informação pode ser consultada pelos utilizadores com maior facilidade (Gao, 2010).

Os *WebSEAD* são tradicionalmente constituídos através de uma relação cliente/servidor de três camadas - cliente, intermédio e armazenamento de dados -, em tudo semelhante às arquiteturas *WebSIG*, descritas por Peng e Tsou (Mileu, 2016). A camada cliente, consiste na sua relação com um servidor, ou seja, o acesso a um *browser* (servidor *URL*), de modo a que seja possível aceder à internet. A camada intermédia, constitui o servidor *Web*, assim como os serviços que permitem efetuar a comunicação entre o cliente e o servidor. *“No contexto do WebSEAD e do WebSIG, em geral, a camada intermédia inclui também um servidor de mapas que permite aceder aos dados geográficos, visualizá-los, bem como efetuar análises e geoprocessamento”* (Mileu, 2016).

2.5. Operacionalização do SEAD

O presente capítulo destina-se à explicação dos procedimentos realizados numa primeira etapa, desde a obtenção ao tratamento da informação. Estes são fulcrais visto

que são a base dos procedimentos aplicados *a posteriori*, nomeadamente, através da execução de métodos estatísticos de apoio à decisão, assim como na implementação de um *WebSEAD*.

2.5.1. Obtenção da Informação

A informação no presente estudo no que concerne a dadores em morte cerebral efetivos foi adquirida através de um levantamento da informação disponível no banco de dados do Sistema de Informação Português para a Transplantação - Organização Portuguesa de Transplantes (listagens da SIOPT), informação disponibilizada ao nível dos hospitais da Rede Nacional de Coordenação de Colheita e Transplantação.

Através dos arquivos dos Autos de Colheita, disponibilizados pelo IPST foi possível proceder à recolha de informação (anonimizada) numa escala mais vasta, o que, posteriormente, possibilitou a geocodificação da residência dos dadores, através do código postal, assim como a recolha de informação referente ao hospital onde este foi dador e a respetiva tipologia do(s) órgão(s) doado(s).

No que concerne a doentes transplantados, através dos arquivos de Comunicações à Administração Central do Sistema de Saúde (ACSS), mais uma vez disponibilizados pelo IPST, obteve-se o código postal, o hospital onde se realizou o transplante e a respetiva tipologia do órgão transplantado.

De modo a adquirir variáveis, que fortalecessem o estudo, recorreu-se ao *website* do INE. Estas são variáveis de cariz populacional, relacionadas com as causas de morte dos dadores em Portugal, assim como com doenças de especificidade associadas a doentes que necessitam de transplante ou foram transplantados. Para o mesmo efeito, procedeu-se a um levantamento do número de camas de cuidados intensivos e de cuidados intermédios dos hospitais pertencentes à Rede Nacional de Coordenação de Colheita e Transplantação, através do documento do Serviço Nacional de Saúde (Rede de Referenciação de Medicina Intensiva, 2016).

2.5.2. Construção de uma base de dados em Microsoft Access

As bases de dados geográficas são fulcrais uma vez que são transacionais, ou seja, podem ser constantemente atualizadas consoante surjam novas informações, ao contrário dos mapas tradicionais, que uma vez impressos permanecem imutáveis ao longo do tempo (Longley *et al.*, 2011).

O primeiro passo deste estágio assentou precisamente na criação e estruturação de uma base de dados que contém os dados utilizados nas análises efetuadas, assim como disponibilizados ao IPST. Após recolha da informação, procedeu-se à sua preparação e construção em *Microsoft Access*.

Toda a informação relativa a dadores de órgãos encontra-se disponibilizada ao nível do município e hospitais do SNS. No que concerne a doentes transplantados a informação é disponibilizada ao município e hospitais com unidade de transplantação. A escala temporal abordada no estudo corresponde ao ano de 2016 e 2017.

Posteriormente, foi gerada informação derivada dos dados iniciais, utilizando um conjunto de inquirições em linguagem de programação SQL¹³, com o intuito de estruturação e produção de informação adicional ao longo dos dois anos.

Tabelas referentes a dadores efetivos:

- Número de dadores por género feminino:

```
SELECT ORIGEM, COUNT (ORIGEM) AS GENERO  
FROM BASEDADOS16  
WHERE GENERO = "F"  
GROUP BY ORIGEM
```

- Número de dadores por género masculino:

¹³ SQL ou *Standart Query Language* é uma linguagem padrão para consultar tabelas ou bancos de dados relacionais (Longley *et al*, 2005).

```
SELECT ORIGEM, COUNT (ORIGEM) AS GENERO
FROM BASEDADOS16
WHERE GENERO = "M"
GROUP BY ORIGEM
```

- Número total de órgãos colhidos:

```
SELECT ORGAO, COUNT(ORGAO) AS N_ORGAO
FROM BASEDADOS16
GROUP BY ORGAO
```

- Causa de morte dos doadores por AVC Isquémico:

```
SELECT ORIGEM, COUNT (ORIGEM) AS CAUSAMORTE
FROM BASEDADOS16
WHERE CAUSAMORTE = "AVC Isquemico"
GROUP BY ORIGEM
```

- Causa de morte dos doadores por AVC Hemorrágico:

```
SELECT ORIGEM, COUNT (ORIGEM) AS CAUSAMORTE
FROM BASEDADOS16
WHERE CAUSAMORTE = "AVC Hemorragico"
GROUP BY ORIGEM
```

- Causa de morte dos doadores por Anoxia:

```
SELECT ORIGEM, COUNT (ORIGEM) AS CAUSAMORTE
FROM BASEDADOS16
WHERE CAUSAMORTE = "ANOXIA"
GROUP BY ORIGEM
```

- Causa de morte dos doadores por TCE-Tiro/Agressão:

```
SELECT ORIGEM, COUNT (ORIGEM) AS CAUSAMORTE
FROM BASEDADOS16
WHERE CAUSAMORTE = "TCE-Tiro/Agressao"
```


GROUP BY ORIGEM

- Causa de morte dos dados por TCE-Acidente de Viação:

```
SELECT ORIGEM, COUNT (ORIGEM) AS CAUSAMORTE
FROM BASEDADOS16
WHERE CAUSAMORTE = "TCE-AcidenteViação"
GROUP BY ORIGEM
```

- Causa de morte dos dados por TCE-Acidente de Trabalho:

```
SELECT ORIGEM, COUNT (ORIGEM) AS CAUSAMORTE
FROM BASEDADOS16
WHERE CAUSAMORTE = " TCE-AcidenteTrabalho"
GROUP BY ORIGEM
```

- Causa de morte dos dados por TCE-Outro:

```
SELECT ORIGEM, COUNT (ORIGEM) AS CAUSAMORTE
FROM BASEDADOS16
WHERE CAUSAMORTE = "TCE-Outro"
GROUP BY ORIGEM
```

- Causa de morte dos dados por Outro:

```
SELECT ORIGEM, COUNT (ORIGEM) AS CAUSAMORTE
FROM BASEDADOS16
WHERE CAUSAMORTE = "Outro"
GROUP BY ORIGEM
```

- Idade média dos dados falecidos:

```
SELECT AVG(IDADE)AS MEDIAIDADE
FROM BaseDados16
```

2.5.3. Geocodificação e Georreferenciação

Por georreferenciação entende-se “o processo de converter informação não-geográfica, *i.e.*, que não tem nenhuma referência geográfica válida que possa ser utilizada em análises espaciais, em informação geográfica” (Hill, 2006).

A geocodificação (*geocoding*) é considerada um método de georreferenciação (Goldberg *et al.*, 2007), através de um processo de conversão de um endereço num ponto referenciado pela latitude e longitude com posição geográfica no espaço Terra (Croner *et al.*, 1996). Porém, tal conceito também pode sofrer transformações consoante a área em que é aplicado, *e.g.*, investigação, indústria ou mundo académico.

Sendo a geocodificação um procedimento automático e dessa forma simples e rápido de executar, para além de recursos básicos como endereços, mapas e conhecimento local da região em estudo, o *geocoding* dispensa qualquer outro recurso. Contudo, para o processo de correção, é indispensável o trabalho manual (Siffel *et al.*, 2006). Tal processo, dadas as suas potencialidades, reverte-se numa componente importante de assimilação de dados que carecem de representação espacial em estudos ligados à saúde (Zimmerman *et al.*, 2008). Este é um método que tem demonstrado um leque de vantagens acrescidas, onde podem e devem ser aproveitadas as suas potencialidades pelos decisores dos diversos níveis de implementação de políticas de saúde (Riner *et al.*, 2004).

Ao longo do estágio, foi disponibilizado pelo IPST uma listagem com os hospitais pertencentes ao SNS, nomeadamente os hospitais com unidade de transplantação e os hospitais da Rede Nacional de Coordenação de Colheita e Transplantação. Dado que a lista de hospitais não detinha uma morada associada, de modo a rentabilizar tempo, recorreu-se a uma pesquisa na aplicação *Google Maps* para obtenção da sua localização, e, posteriormente, através do método de geocodificação de pesquisa única, obter as coordenadas de cada hospital. À lista de hospitais em tabela, adicionou-se um campo de longitude e um campo de latitude com as coordenadas obtidas. No *software ArcGis*

versão 10.6.1, foi possível importar tal informação, assim como convertê-la em estrutura vetorial (pontos).

No que concerne a dados efetivos procedeu-se a um levantamento das moradas, mais especificamente os códigos postais. Dado que este é um levantamento exaustivo de informação existe sempre uma pequena margem para erro. Na generalidade, um dos entraves que se coloca está sobretudo relacionado com a existência de erros associados. A precisão de tal processo depende não só da qualidade da base de dados que é utilizada para gerar a georreferenciação, mas também da integridade e exatidão dos endereços fornecidos (Siffel *et al.*, 2006).

Dos 268 dados falecidos efetivos¹⁴ introduzidos na base de dados, referente ao ano de 2016, foi possível georreferenciar 226, o que se traduz numa taxa de sucesso de 85%. Quanto ao ano de 2017, dos 327 dados efetivos, foi possível georreferenciar 286 códigos postais, o que se traduz numa taxa de sucesso de 87%.

Quanto aos doentes transplantados, recolhidos os códigos postais, deu-se início ao processo de geocodificação. Dos 622 utentes transplantados registados na base de dados referente ao ano de 2016, foi possível georreferenciar 561, o que se traduz a uma taxa de sucesso de 90,1%. Relativamente ao ano de 2017, dos 926 utentes, georreferenciaram-se 797 casos, ou seja, uma taxa de sucesso de 86%. Harries (1999) afirma que embora não haja um padrão mínimo para a geocodificação é possível produzir e distribuir mapas com base numa taxa de sucesso a partir de 25%.

¹⁴ “Um dador elegível com consentimento, em quem foi feita uma incisão cirúrgica com a intenção de colher órgãos para transplantação e de quem se colheu, pelo menos, um órgão para transplantação.” em: Guia para a qualidade e segurança dos órgãos para transplantação, 2013.

Capítulo 3. Métodos de análise e classificação de dados

A estatística espacial compreende um conjunto de técnicas para descrever e modelar dados espaciais (Scott e Janikas, 2010). No contexto da pesquisa em saúde pública, os métodos estatísticos utilizados em ambiente SIG e as técnicas de detecção remota¹⁵ são essenciais para a formulação de hipóteses. Estes, devem ser utilizados na exploração e quantificação da significância estatística de tendências observadas na localização e distribuição espacial (Blossom, 2011).

O presente capítulo destina-se a descrever os métodos estatísticos utilizados na análise de padrões efetuada ao longo do estudo, tendo estes sido selecionados com o intuito de analisar as características dos dados relativamente à sua distribuição espacial.

O ponto de partida tem por método o Índice Global e Local de Moran, através da ferramenta de análise de *clusters*, com o intuito de confirmar a autocorrelação espacial do número de doentes transplantados e, determinar quais as aglomerações estatisticamente significativas, sejam compostas por valores semelhantes elevados ou baixos. Estimada a autocorrelação espacial, surge o interesse em compreender o grau de aglomeração dos dados efetivos. Para que seja possível o cálculo desse índice, foi necessária a elaboração de uma matriz de pesos espaciais, i.e., noção de proximidade entre as observações.

De modo a perspetivar a significância dos valores de *clustering* nos municípios estatisticamente significativos face ao número de doentes transplantados nas diversas tipologias de órgãos, recorreu-se ao método de análise de *Hot Spot*.

Para o estudo de padrões espaciais e exploração de relações existentes, aplicaram-se os métodos de análise de Regressão Linear Múltipla (RLM), Árvores de Regressão e Classificação (*Classification and Regression Trees* [CRT]) e Redes Neurais Artificiais (RNA). Quanto ao método RLM, do padrão observado dos doentes transplantados,

¹⁵ Considera-se “o conjunto de técnicas que possibilitam a obtenção de informações sobre elementos na superfície terrestre, realizado por sensores distantes, ou remotos” DGT, 2018.

foram selecionadas 15 variáveis que representam as causas de óbitos dos doadores falecidos efetivos (c.f capítulo 6). O que se pretende com a aplicação deste método é evidenciar as áreas (municípios) com potencial de doação. O mesmo processo foi realizado para o padrão observado de doentes transplantados, através da seleção de 11 variáveis (c.f capítulo 6), que representam patologias que poderão levar à falência de órgãos e à necessidade de transplante e, como tal, poderão ter influência na explicação do número de doentes transplantados. O pressuposto neste caso é demonstrar quais as áreas onde o número de doentes a necessitar de um transplante pode vir a aumentar. Por fim, com o intuito de confrontar métodos, optou-se pela utilização de dois métodos atualmente em voga, as CRT e as RNA, para perspetivar qual o conjunto de variáveis que melhor justificam a localização dos doentes transplantados.

De modo a construir a ponte entre os métodos utilizados e os objetivos concretos do estudo, apresenta-se o Quadro 3.1 onde é possível verificar os métodos estatísticos a aplicar, sintetizando, a sequência dos procedimentos e o seu propósito.

Quadro 3.1. Esquema da aplicação dos métodos de análise espacial

Método Estatístico	Propósito
Índice Global de Moran's	Perspetivar se existe significância estatística nas observações, assim como a existência de autocorrelação face ao número de doentes transplantados nas várias tipologias de órgãos.
Índice Local de Moran's	Identificar os municípios de concentração de doadores efetivos, através da análise baseada na relação espacial entre as observações, caso sejam constituídos por valores semelhantes, elevados ou baixos tendo em conta a matriz de vizinhança.
Análise de Hot Spot (Getis-Ord Gi*)	Detetar a significância dos valores de <i>clustering</i> nos municípios estatisticamente significativos face ao número de doentes transplantados nas diversas tipologias de órgãos.
Regressão Linear Múltipla	Identificar os fatores que influenciam a localização geográfica dos doadores efetivos e doentes transplantados, de modo a apurar as áreas com potencial de doação e as áreas onde o número de doentes pode vir a aumentar.
Árvores de regressão e classificação	Face às variáveis independentes predefinidas qual a estrutura de árvore e variáveis que melhor justificam a localização dos doentes transplantados.
Redes Neurais Artificiais	Face às variáveis independentes predefinidas qual a rede e variáveis que melhor justificam a localização dos doentes transplantados.

3.1. Índice de Moran

O Índice de Moran é uma estatística global, que não permite apreciar a estrutura local da autocorrelação espacial. No entanto é possível questionar quais as regiões que mais contribuem para a autocorrelação espacial global, se existem aglomerados espaciais locais de valores altos e baixos, ou até que ponto tal avaliação global se desvia do padrão de auto correlação espacial positiva (Gallo e Ertur, 2000).

Em estatística, a autocorrelação *"corresponde a uma medida que quantifica quanto o valor de uma variável aleatória é capaz de influenciar as unidades espaciais vizinhas, i.e., um valor alto condiciona também os valores altos dos seus vizinhos? Os índices de auto correlação espacial testam o grau de dependência espacial dos territórios que são vizinhos (Rocha, 2012)"*.

O índice de Moran pode ser definido como um indicador que mede o grau de concentração ou dispersão da distribuição espacial de um conjunto de dados, numa escala global, permitindo identificar se o padrão espacial apresenta uma autocorrelação positiva ou negativa (Lee e Wong, 2001). O Índice Global de Moran é dado pela seguinte expressão:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} Z_i Z_j}{\sum_{i=1}^n Z_i^2} \quad (3.1)$$

Onde Z_i é o desvio de um atributo do elemento i relativamente à sua média, $w_{i,j}$ é a ponderação espacial entre i e j , n corresponde ao número total de elementos e S_0 é a agregação de todos os pesos espaciais.

De uma visão geral, o índice de Moran *"corresponde a um teste cuja hipótese nula é a independência espacial, sendo que neste caso o seu valor seria zero. Valores positivos (entre 0 e +1) indicam uma correlação direta e valores negativos, (entre 0 e -1) uma correlação inversa (Rocha, 2012)"*. O valor deste índice varia entre -1 e 1, e é (1) positivo

quando áreas próximas tendem a possuir atributos semelhantes - padrão concentrado, (2) negativo quando subsistem valores altos próximos de valores baixos - padrão disperso, e (3) sensivelmente zero quando os atributos estão distribuídos forma aleatória no espaço (Longley *et al.*, 2011).

A análise de autocorrelação espacial local é realizada através de duas ferramentas: inicialmente através do diagrama de dispersão de Moran, e numa segunda instância, através de Indicadores Locais de Associação Espacial (LISA) (Anselin, 1995).

Denominam-se por LISA os indicadores que tem por finalidade quantificar o grau de associação espacial, a que cada localização do conjunto da amostra se encontra sujeita em função de um modelo de vizinhança pré-estabelecido (Rocha, 2012). Tais indicadores *“distinguem a relação entre o valor de um atributo relativo a uma entidade, com os valores correspondentes às entidades na sua vizinhança e consegue identificar a extensão da concentração espacial de valores semelhantes ao redor de cada observação”* (Anselin, 1995). Um dos métodos LISA mais divulgados é o Índice Local de Moran (I_i), expresso pela seguinte equação:

$$I_i = \frac{A_i - \bar{X}}{S^2} \sum_{j=1, j \neq i}^i \mathcal{W}_{i,j} (X_j - \bar{X}) \quad (3.2)$$

Onde A_i é o atributo do elemento i , $\mathcal{W}_{i,j}$, representa o peso espacial (grau de conexão) entre as duas *features*, \bar{X} é a média do atributo correspondente e S^2 é calculado por:

$$S^2 = \frac{\sum_{j=1, j \neq i}^i \mathcal{W}_{i,j} (X_j - \bar{X})}{n-1} \quad (3.3)$$

Onde n equivale ao número total de observações.

Anselin (1995) aponta a existência de uma simetria entre o valor da autocorrelação global e o valor da autocorrelação local. Através dos métodos LISA, é possível identificar

clusters, i.e., áreas onde valores elevados encontram-se rodeados por valores altos e vice-versa. Além disso, também é possível reconhecer irregularidades no padrão (*outliers*), i.e., áreas onde os valores altos estão rodeados por valores baixos e valores baixos rodeados por valores altos (Longley *et al.*, 2005; Waller e Gotway, 2004; Anselin *et al.*, 2010).

3.2. Análise de Hot Spot (Getis-Ord G_i^*)

A estatística de G_i^* originalmente desenvolvida por Getis Ord, é utilizada no estudo das evidências de padrões espaciais identificáveis. Muito semelhante ao Índice de Moran, esta é uma estatística global dado que o grau de correlação espacial é considerado.

Face à sua generalidade, os padrões locais provavelmente serão neutralizados quando deparados com uma vasta área, i.e., tornar-se-ão indetetáveis. No entanto, dado que o nível de dependência espacial pode alternar significativamente ao longo do espaço, sugere a capacidade de identificar heterogeneidade espacial (Mitchell, 2005). O Índice Global de Moran de certa forma acaba por compensar algumas limitações deste modelo, no entanto, não é apto no que concerne à discriminação e identificação de valores *hot spots* (pontos quentes) e valores *cold spots* (pontos frios) no território. O Índice de Getis-Ord G_i^* é adequado quando o pressuposto é a localização de regiões a uma escala global, assim como a diferenciação de *clusters* - valores altos e baixos entre as observações locais (Mitra, 2009).

O Índice de Getis-Ord G_i^* é dado pela seguinte equação:

$$G_i^*(f) = \frac{\sum_j W_{ij}(f) X_j}{\sum_j X_j} \quad (3.4)$$

Onde (d) representa a estatística de G_i^* e inclui o valor na localização dentro da medida de agrupamento e W_{ij} é a matriz normalizada que determina a vizinhança entre as áreas.

Este índice fornece valores de *z-score* e *p-value*, (Figura 3.1) tendo por utilidade a averiguação de significância estatística nas observações, caso contrário apura-se a hipótese nula (Mitra, 2009). O valor de *p-value* é uma probabilidade, i.e., quando o valor é muito pequeno expressa a elevada improbabilidade de o padrão espacial observado ser resultado de um processo aleatório. Quanto ao *z-score* este traduz-se num desvio padrão, e tanto o *p-value* como o *z-score* estão associados a uma distribuição normal (Mitchell, 2005).

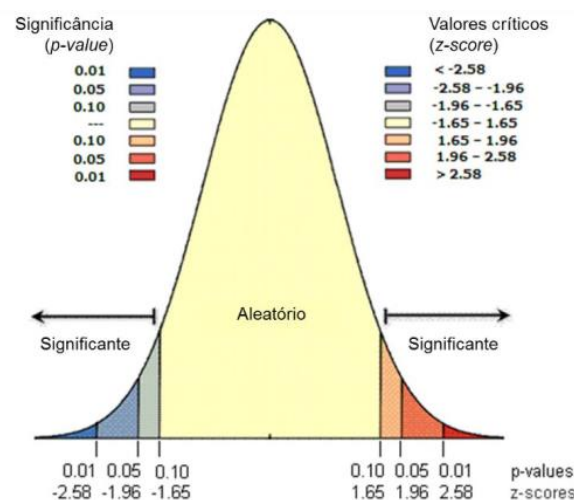


Figura 3.1. Distribuição normal dos *p-values* e dos *z-scores* (ArcGis help).

Antes de qualquer tipo de cálculo, é imprescindível definir o método para concetualizar a relação espacial entre as observações (*features*). No caso do presente estudo, optou-se pela distância inversa (*inverse distance*), i.e., as observações vizinhas próximas têm uma maior influência sobre os cálculos de uma *target feature* do que as observações que se encontram mais distantes.

Depois de calculados os resultados, estes podem variar entre valores positivos e negativos. Quando o *z-score* for positivo, assim como quanto mais elevado for o seu valor, maior será o aglomerado de *hot spot*, porém, quando o valor de *z-score* for negativo, menor será o seu valor, logo maior a aglomeração de *cold spot*.

3.3. Regressão Linear Múltipla

A análise de regressão permite modelar, examinar e explorar as relações espaciais, visando o auxílio e explicação de fatores por detrás dos padrões espaciais observados (Encalada, 2016).

O método dos Quadrados Mínimos (*Ordinary Least Squares*) é o ponto de partida adequado para todas as análises de regressão espacial. Dado que fornece um modelo global da variável ou processo que se está a tentar compreender ou prever, o método gera uma equação única de regressão simples.

Um dos primeiros aspetos a considerar quando se esquematiza um modelo de regressão linear é a seleção dos fatores explicativos, i.e., as variáveis a utilizar. Efetuada a escolha das variáveis, tal técnica permite medir o grau de dependência que a variável dependente (Y) apresenta em relação à variável independente (X) (Ebdon, 1985). Cada variável independente está associada a um coeficiente de regressão, o qual descreve a força e o sinal da relação dessa variável com a variável dependente.

A melhoria do desempenho de um modelo de regressão linear, não se encontra apenas relacionado com a correlação da variável dependente com a variável independente, mas também com a conexão dessa variável (independente) com as demais variáveis independentes do modelo. Em situações onde essas dependências forem fortes, declara-se a existência da presença de multicolinearidade. Esta ocorre quando, e.g., duas variáveis medem aproximadamente o mesmo fenómeno. Um das possíveis formas de deteção de multicolinearidade é através do valor de tolerância ou o seu inverso, designado por Fator de Inflação da Variância (VIF) e é dado pela seguinte equação:

$$VIF = \frac{1}{1 - R_j^2} \quad (3.5)$$

Onde R_j resulta da regressão de x_j com as outras variáveis.

O VIF é uma medida de grau, em que cada variável independente é explicada pelas demais variáveis, i.e., quanto maior o valor de VIF mais severa se revelará a multicolinearidade (Rocha, 2012). Caso ocorra colinearidade algumas das variáveis explicativas devem ser retiradas do estudo. Cada pesquisador deve determinar o grau de colinearidade aceitável perante o assunto estudado, uma vez que a maioria dos padrões ou limites recomendados ainda permitem colinearidade substancial. Hair, *et al.* (1995) sugerem que quando o tamanho das amostras é menor há a possibilidade de ser-se mais restrito devido ao aumento no erro padrão, sendo possível optar por um limiar de corte de VIF com tolerância máxima de valor 10. Logo, apenas as variáveis com correlações de valor VIF inferior a 10 devem ser incluídas num modelo de regressão, para que não haja risco de ocorrência de multicolinearidade elevado.

No seguimento do processo, dá-se o cálculo da Regressão Geográfica Ponderada (RGP) (*Geographically Weighted Regression*). Esta é aplicada às variáveis explicativas que se denotam como aquelas que auxiliam a predizer o comportamento da variável dependente que se pretende modelar. Assim, tem-se a equação:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 \dots + \beta_n X_n \quad (3.6)$$

Onde Y é a variável dependente, α o intercepto do eixo Y , X as variáveis explicativas e β os coeficientes de regressão.

O objetivo primordial é o fornecimento de estatísticas importantes e confiáveis, com o intuito de examinar e estimar relacionamentos lineares, e sobretudo localizar em antecipação alterações conjeturáveis. A RGP disponibiliza um modelo local ou o processo que se pretende modelar e ajusta uma equação de regressão a cada fator. Estas podem ser negativas (abaixo do espectável) ou positivas (acima do espectável).

3.3.1. Validação

Uma medida da qualidade de ajuste é o R^2 (coeficiente de determinação), que em mínimos quadrados ordinários relativos a um intercepto variam entre 0 e 1. Um R^2 baixo implica que o modelo não se encaixa bem nos dados. Contudo um R^2 próximo de 1 faz não garante que o modelo se encaixe bem nos dados: como Anscombe (1973) demonstrou, um R^2 elevado pode ocorrer na presença de erros de especificação da forma funcional de um relacionamento ou na presença de *outliers* que distorcem a relação verdadeira.

Um problema com o R^2 como uma medida de validade do modelo é que ele pode sempre ser aumentado adicionando mais variáveis ao modelo, exceto no caso improvável de que as variáveis adicionais sejam exatamente não correlacionadas com a variável dependente na amostra de dados que está a ser utilizada. Este problema pode ser evitado fazendo-se um teste-F da significância ou através do VIF, ou simplesmente analisando o R^2 ajustado.

3.3.2. Vantagens e desvantagens

Há duas vantagens principais na análise de dados usando um modelo de regressão múltipla. A primeira é a capacidade de determinar a influência relativa de uma ou mais variáveis preditoras na variável dependente. A segunda vantagem é a capacidade de identificar *outliers* ou anomalias.

A sua principal desvantagem é que muitos fenómenos do mundo real simplesmente não correspondem às suposições de um modelo linear. Nesses casos é difícil, ou mesmo impossível, produzir resultados úteis através de uma regressão linear. Qualquer outra desvantagem resultante da utilização deste modelo advém das características dos dados utilizados, e.g., dados incompletos ou concluir erradamente que uma correlação é uma causa.

3.4. Árvores de classificação e regressão

As árvores de classificação e regressão (Classification and Regression Trees [CRT]), constituem um método recursivo de particionamento que elabora estruturas em árvore com o intuito de prever variáveis dependentes contínuas (regressão). O algoritmo clássico de CRT, criado em 1884 (Breiman et al., 1984; Ripley, 1996), procura particionar recursivamente a variável de resposta em subconjuntos binários cada vez mais puros através de critérios de divisão e de paragem (Venables e Ripley 2002). A árvore pode crescer de forma a ajustar-se exatamente aos dados de treino. Este método tem várias vantagens: i) pode lidar com qualquer combinação de dados categóricos e contínuos na regressão; ii) os resultados da CRT são apresentados por um conjunto de divisões lógicas *se* → *então* que permitem a exata predição e classificação dos casos, e por conseguinte, são fáceis de interpretar; e iii) tem a capacidade de capturar a relação hierárquica e não-linear entre as variáveis de predição (Death e Fabricius, 2000).

Existem inúmeros algoritmos para prever variáveis contínuas com base num conjunto de preditores contínuos e/ou categóricos. Os problemas tipo regressão são geralmente aqueles onde se deve tentar prever os valores de uma variável contínua a partir de uma ou mais variáveis preditivas contínuas e/ou categóricas. Usando uma regressão linear múltipla, ir-se-ia determinar uma equação linear para cada uma das variáveis que podem ser usadas para calcular a variável dependente. Há muitos procedimentos analíticos diferentes para ajustar os modelos lineares, e.g. regressão, vários tipos de modelos não-lineares, ou modelos não-lineares completamente personalizadas, onde se pode incluir uma equação arbitrária que contendo parâmetros para serem estimados.

3.4.1. Regras de divisão (*medidas de impureza dos nós*)

Começando a partir da raiz e recorrendo aos dados de teste, a amostra vai sendo dividida e catalogada de acordo com uma regra de divisão binária. Se os dados separados pertencem à mesma classe, são combinados de forma a constituir uma folha. Se, pelo contrário, existem ocorrências de diferentes classes, o processo de divisão

contínua. Existem três algoritmos principais de divisão, *i.e.*, o de entropia, o rácio de ganho e o de Gini. No primeiro caso temos

$$Ent = \sum_{j=1}^k \frac{freq(C_j, S)}{|S|} \times \log_2 \left(\frac{freq(C_j, S)}{|S|} \right) \quad (3.7)$$

Com $|S|$ correspondendo ao número de ocorrências no grupo S e C_j ao número de ocorrências da classe j no grupo S .

O ganho de uma única classificação X é definido como a entropia depois dessa mesma classificação X . Assim temos

$$Ent_X(S) = \frac{|S_1|}{S} \times Ent(S_1) \quad (3.8)$$

O ganho (X) traduz-se por $Ganho(X) = Entropia(S) - Entropia_X(S)$ e testa a maximização do ganho de informação. O algoritmo da entropia é muito dado à sobredivisão, uma vez que, potencialmente, cada divisão pode contribuir para o ganho de informação. O rácio de ganho tenta sublevar esta *bias* potencial através de um processo de normalização. Se definirmos a informação da divisão de (X) como

$$Inf_div(X) = - \sum_{i=1}^n \frac{|S_i|}{S} \times \log_2 \left(\frac{|S_i|}{S} \right) \quad (3.9)$$

Que representa o potencial de informação gerado ao dividir S em n subconjuntos. Então, o ganho de informação mede os dados como

$$Rac_gan(X) = Gan(X) / Inf_div(X) \quad (3.10)$$

Onde o rácio de ganho - $Rac_gan(X)$ - tende a maximizar a razão (3.10).

A regra de divisão de Gini é uma medida de impureza de um dado inter-nó, que assume o seu valor máximo quando todas as ocorrências estão equitativamente distribuídas por todas as classes. De uma forma geral, esta regra tenta encontrar a maior categoria homogénea dentro do conjunto de dados e isolá-la das restantes ocorrências (Zambon et al., 2006).

$$Gini(S) = \sum_i freq(C_j, S) \times (1 - freq(C_j, S)) \quad (3.11)$$

De (3.11) retira-se uma medida relativa de quão bons são os dados de treino. Um valor positivo significa que a classe adquiriu ocorrências de outra classe, enquanto um valor negativo significa que a classe perdeu ocorrências para outra classe (c.f análise de erro nos mapas de uso do solo e validação dos modelos, neste capítulo). O que é importante é a quantidade de classes que se desviam de zero e a magnitude desse desvio. A equação utilizada para determinar o grafo final é

$$\left| \sum_{i=1}^k (C_i p_{i+}) \right| = \left| \sum_{i=1}^k (C_i p_{i-}) \right| \quad (3.12)$$

Onde C_j é o número inicial de ocorrências observadas para a classe i , e k o número total de classes. P_{i+} é a proporção (positiva) de ocorrências da classe i sobreclassificadas; no fim $P_{i+} > 0$. P_{i-} é a proporção (negativa) de ocorrências da classe i subclassificadas; no fim $P_{i-} < 0$. A equação (3.12) significa que o número total de ocorrências sobreclassificadas para todas as classes (ganho total) deve ser igual ao número total de ocorrências subclassificadas em todas as classes (perda total).

$$p_i = \frac{A_i - C_i}{C_i} \quad (3.13)$$

Em que A_i é o número de ocorrências realmente classificadas da classe i ; Se $A_i - C_i < 0$, p_i torna-se p_{i-} , se $A_i - C_i > 0$, p_i torna-se p_{i+} .

3.4.2. Cálculo da exatidão na regressão

Nos problemas de regressão (variáveis dependentes contínuas) podem ser aplicadas as mesmas três medidas de exatidão. No primeiro - estimação de ressubstituição - caso passamos a ter

$$R(d) = \frac{1}{N} \sum_{i=1}^N (y_i - d(x_i))^2 \quad (3.14)$$

Onde a amostra de aprendizagem z consiste de (x_i, y_i) , $i = 1, 2, \dots, N$. Tal como na vertente de classificação, a estimação de ressubstituição é calculada utilizando os mesmos dados que serviram para construir o classificador (d) .

O teste da amostra fica como

$$R^{ta}(d) = \frac{1}{N_2} \sum_{(x_i, y_i) \in z_2} (y_i - d(x_i))^2 \quad (3.15)$$

Onde z_2 é a subamostra que não é utilizada para criar o classificador.

Por fim, a validação cruzada v-pastas passa a apresentar a forma

$$R^{vc}(d) = \frac{1}{N_v} \sum_v \sum_{(x_n, y_n) \in z_v} (y_n - d^{(v)}(x_n))^2 \quad (3.16)$$

Onde $d^{(v)}(x)$ é calculado a partir da subamostra $z - z_v$.

3.4.3. Vantagens e desvantagens

Há um grande número de métodos que um analista pode escolher para tratar problemas de classificação ou de regressão. As técnicas de classificação em árvore, quando funcionam e produzem predições exatas ou classificações, baseadas nalgumas

condições lógicas do tipo *se* \rightarrow *então* , têm uma série de vantagens sobre muitas dessas técnicas alternativas.

Na maioria dos casos, a interpretação dos resultados resumidos numa árvore é muito simples. Esta simplicidade é útil não somente para fins de uma classificação rápida de novas observações (é muito mais fácil de avaliar apenas uma ou duas condições lógicas, do que ao calcular a pontuação de classificação para cada grupo de possível, ou prever valores, com base em todos os preditores e usando possivelmente algumas equações de modelos complexos não-lineares), mas também muitas vezes pode providenciar um modelo muito mais simples para explicar por que é que as observações são classificadas ou preditas de uma forma particular.

Os métodos em árvore são não-paramétricos e não-lineares. Os resultados finais da utilização de métodos de árvore de classificação ou regressão podem ser resumidos numa série de condições (normalmente poucas) lógicas (nós da árvore). Portanto, não há nenhum pressuposto implícito que as relações subjacentes entre as variáveis de predição e a variável dependente sejam lineares, sigam algumas funções específicas de ligação não-lineares, ou que até são monótonas na natureza.

Por exemplo, a árvore poderia revelar várias divisões baseadas na mesma variável, revelando uma relação não monótona entre as variáveis. Assim, os métodos em árvore são particularmente adequados para tarefas de pesquisa de dados (*data mining*), onde há muitas vezes pouco conhecimento *a priori* e não há nenhum conjunto coerente de teorias ou predições sobre quais variáveis são relacionadas e como. Nesses tipos de análises de dados, os métodos em árvore podem revelar muitas vezes, relações simples entre apenas algumas variáveis que poderiam facilmente ter passado despercebidos usando outras técnicas analíticas.

Embora sendo um dos métodos não paramétricos que melhor se adapta à aproximação de funções gerais multivariadas, o método CRT de (Breiman et al., 1984) apresenta certas restrições que limitam fortemente a sua eficiência. A utilização da metodologia de atribuir à função de aproximação valores constantes nas sub-regiões do espaço de

predição determinadas pelas divisões, conduz inevitavelmente à obtenção de funções de aproximação fortemente descontínuas nas fronteiras dessas sub-regiões e é uma das causas principais da ineficiência deste método quando aplicado a situações em que a função a aproximar é contínua.

Outro problema do método CART é a extrema dificuldade que ele tem em aproximar funções lineares ou aditivas com mais do que uma variável; de um modo geral, este método aproxima com dificuldade funções nas quais as interações existentes envolvem um pequeno número de variáveis. A necessidade de colmatar estas (e outras) limitações motivou o desenvolvimento de novas metodologias.

3.5. Redes Neurais Artificiais

Parte da investigação em Redes Neurais Artificiais (RNA) encontra-se inspirada e influenciada pelos sistemas nervosos dos seres vivos, com particularidade no do ser humano. O cérebro é um sistema de processamento de informação, assemelhando-se a um computador, uma vez que é altamente complexo, não-linear e paralelo (Haykin, 1999). Segundo o Investigador Teuvo Kohonen (Kohonen, 2014) da Universidade de Helsinki, uma RNA é *“uma rede massivamente paralela de elementos interconectados e organizações hierárquicas, que estão preparadas para conjugar objetos do mundo real, do mesmo modo que um sistema nervoso biológico faz”*. Por outras palavras, as RNA são sistemas adaptativos, com o intuito de criar circuitos com comportamentos similares ao de um circuito cerebral (neurónios humanos), daí o termo rede neuronal. Estes sistemas assemelham-se ao cérebro humano no processamento de informação em dois aspetos:

- 1) O conhecimento é adquirido pela rede através de um processo de aprendizagem;
- 2) O conhecimento adquirido é armazenado pelas forças de ligação interneuronais.

As RNA têm como implícito a construção de um modelo composto por um grande número de unidades de processamento muito simples, designados por neurónios, com um grande número de conexões entre eles. Estas possuem uma propensão natural para armazenar conhecimento empírico e torná-lo acessível ao utilizador. Deste modo, uma

RNA é composta por diversas unidades de processamento, i.e., os nós e as conexões que ligam os nós. O processamento básico de informação da rede ocorre nos neurônios, sendo a informação transmitida através de conexões denominadas *sinapses* ou *pesos sinápticos* (Santos, *et al.*, 2005).

3.5.1. Perceptrão multi-camada (MLP)

Desenvolvido originalmente por David E. Rumelhart e James L. McClelland (1986), a criação do perceptrão levou ao desenvolvimento de uma das primeiras redes neurais de perceptrão multicamada (*Multi-Layer Perceptron* [MLP]) (Rumelhart, Hinton e Williams, 1986), que atualmente constituem uma das estruturas de RNA mais utilizadas (Figura 3.2).

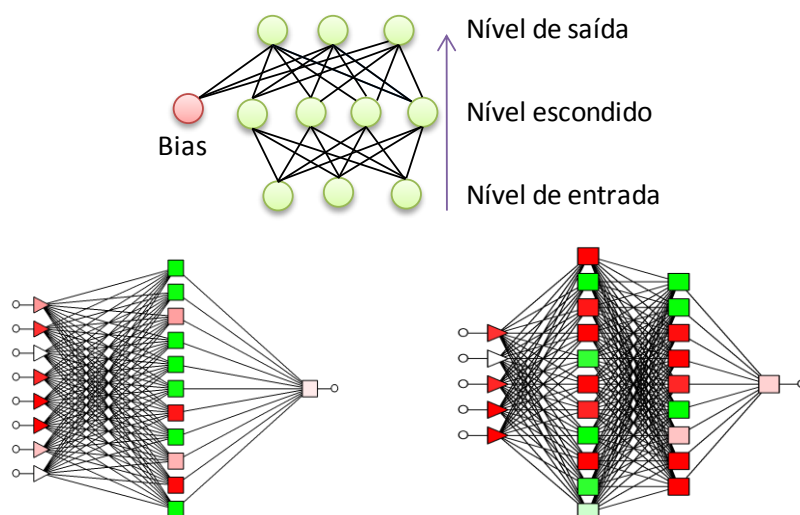


Figura 3.2. Esquema de uma MLP completamente conectada. Note-se que o nível escondido e o de saída possuem bias.

Nestas redes podem-se encontrar três tipos de células: células de entrada, cujo valor de saída é fixado externamente; células de saída, que devolvem a saída da rede; e finalmente, as células intermédias, que não são de entrada nem de saída (Gallant, 1993). Ou seja, temos três níveis (Lippmann, 1987): (1) nível de entrada, (2) nível escondido e (3) nível de saída, e possui a vantagem de poder identificar relacionamentos de natureza não-linear.

As redes com uma ou mais camadas intermédias são chamadas de redes multicamada. Está provado que uma RNA com uma camada intermédia pode representar uma qualquer função contínua, e que com duas camadas intermédias até funções descontínuas podem ser representadas (Russel e Norving, 1995).

Os algoritmos que compõem as RNA calculam os pesos das variáveis e dos nós dos níveis de entrada, escondido(s) e saída recorrendo à introdução de dados iniciais (de entrada) através de um processo designado de alimentação para a frente, que os propaga através dos níveis escondido(s) e de saída. Os sinais propagam-se de nó para nó, sendo modificados pelo peso associado a cada ligação.

O nó (neurão) recetor procede à soma ponderada dos valores de todos os nós do nível anterior que lhe estão conectados. A saída deste nó (output) é então calculada em função dos valores ponderados de entrada, sendo designada de função de ativação (Rocha *et al.*, 2004). Deste modo, os dados movem-se para a frente, de nó para nó, com a ocorrência de múltiplos somatórios ponderados, até atingirem o nível de saída. Para cada camada neuronal da rede MLP, existe um termo *bias*¹⁶. Uma *bias* é um neurão em que a função de ativação é permanentemente fixada em 1. Tal como os outros neurões, a *bias* liga-se aos neurões da camada seguinte através de um peso, muitas vezes designado de limiar (*threshold*). Deste modo, a rede acaba por ser uma interpretação simples do modelo de entrada-saída, constituindo os pesos e os limiares os parâmetros livres, i.e. ajustáveis, do modelo.

Acharya *et al.* (2003) identificam três decisões importantes no processo de construção de uma rede neuronal: (1) a topologia da rede; (2) o algoritmo de aprendizagem e (3) a função de ativação.

¹⁶ Bias (ou função bias) de um estimador é a diferença entre o valor expectável desse estimador e o verdadeiro valor do parâmetro que está a ser estimado. Um estimador ou uma regra de decisão com uma bias de zero diz-se não-biasado. Caso contrário é biasado.

3.5.2. O Algoritmo de Retropropagação (avaliação do erro)

Atualmente, o algoritmo de retropropagação é a técnica de aprendizagem supervisionada mais utilizada em redes neuronais unidirecionais multicamada (Gallant, 1993). A maioria dos programas de RNA utilizam a implementação deste algoritmo, na sua forma original (através do gradiente), ou numa forma modificada com o intuito de melhorar o desempenho da regra. Baseado na regra delta apresentado por Widrow e Hoff, este também pode ser denominado como regra delta generalizada, para redes diretas com mais de duas camadas.

O algoritmo de retropropagação consiste numa propagação positiva do sinal funcional e numa retropropagação do erro (MacKay, 1991), o designado um passo para a frente e outro para trás. O passo para frente denomina-se de propagação e nele os valores provenientes dos neurónios de entrada são aplicados aos neurónios ocultos e, posteriormente, as saídas são aplicadas como entradas aos neurónios da camada final, obtendo a resposta da rede (Haykin, 1999). No passo para atrás, ajustam-se os pesos sinápticos através do cálculo do erro realizado na camada de saída, e os pesos sinápticos entre as camadas antecessoras são ajustados de acordo com uma regra de correção de erro. Por outras palavras, durante o treino com o algoritmo de retropropagação a rede opera numa sequência de duas fases (Haykin, 1999), onde:

- 1) Um padrão é exposto à camada de entrada da rede. O resultado da atividade flui através da rede, camada por camada, até que a camada de saída produza uma resposta;
- 2) A saída obtida é comparada com a saída pretendida. O algoritmo tem por função o cálculo do erro na saída da rede e a retropropagação pela rede, alternando os pesos de forma a minimizar o erro da próxima saída.

Segundo o proposto por Lippmann (1987), o passo inicial para o funcionamento do algoritmo de retropropagação, corresponde ao processo de atribuição dos valores iniciais aos pesos sinápticos, i.e., os valores iniciais dos pesos associados às ligações entre *nodos*. Estes deverão ser pequenos e gerados aleatoriamente e, em princípio,

pertencendo ao intervalo $\{-2/z, 2/z\}$, para uma célula com z entradas (Gallant, 1993). Por vezes adotam-se parâmetros ainda mais restritivos, do tipo $\{-0,1; 0,1\}$. Dado o passo inicial, parte-se para o treino da rede. Este inicia com a seleção de um caso para treino, seguindo-se o cálculo do gradiente e o ajuste dos pesos. A interação só é dada como terminada quando todos os casos disponíveis tiverem sido considerados, e o término do processo só ocorre quando as mudanças nos pesos e no Erro quadrático Médio (EQM) forem insignificantes. Sendo o treino supervisionado torna-se necessário apresentar os padrões de entrada, previamente conhecidos, juntamente com os de saída.

O cálculo dos valores de saída é efetuado através da aplicação do campo local induzido (v) (Haykin, 1999), *i.e.*, a uma função de ativação. Formalmente, e através dos conceitos matemáticos de Freeman e Skapura (1992), o campo local induzido para o neurónio j na camada iteração n é obtido por:

$$v_j^l(n) = \sum_{i=1}^{r+1} W_{ji}^l(n) Y_i^{l-1}(n) \quad (3.17)$$

Onde v é o sinal de saída do neurónio na camada anterior, W é o peso sináptico do neurónio da camada, que é alimentado pelo neurónio da camada e n o número de neurónios na camada anterior.

3.5.3. Vantagens e desvantagens

Rocha (2012) apresenta uma lista onde faz referência a algumas vantagens na utilização de RNA:

1. Nalguns casos permitir análises com resultados superiores aos conseguidos com técnicas estatísticas convencionais;
2. Não necessitam de conhecimentos de especialistas para tomar decisões, pois baseiam-se unicamente nos dados de treino que lhes são fornecidos;

3. O tempo necessário para se implementar uma rede é menor que o utilizado para a construção de um sistema heurístico equivalente;
4. Como as unidades da rede operam em paralelo, a destruição ou defeito num de seus nós não torna a rede inoperante;
5. As redes conseguem separar o ruído (variações aleatórias, presentes nos dados reais, adicionadas aos valores originais) da informação relevante, tendo inclusive sido utilizadas como filtros de dados;
6. Se as condições de trabalho se alterarem pode-se fazer uso da capacidade das RNA para adaptar-se a essas mudanças, voltando a treinar o sistema com novos exemplos.

E como principais desvantagens aponta:

1. Podem chegar a conclusões que contrariem as regras e teorias estabelecidas;
2. O treino de uma rede, dependendo de sua aplicação, pode ser demorado;
3. Pode ser difícil interpretar o seu funcionamento. Trata-se de um tipo de funcionamento denominado de caixa negra;
4. Necessita de um grande volume de dados de treino. A carência deste tipo de dados pode tornar a rede inaplicável.

Neste sentido, e mais concretamente no contexto da Geografia, dever-se-á ter uma posição bastante crítica e atenta ao desenvolvimento destes modelos computacionais. Contudo, não se deverá adotar posições pessimistas e fechadas na utilização destes métodos. De facto, a geografia está hoje no início da sua mais recente revolução que é a era dos *Big Data*. A esperança é que no curso dos próximos 50 anos a geografia interligará todas as suas subdisciplinas num processo de modelação baseado numa estrutura cibernética, capaz de fornecer um entendimento racional e científico dos sistemas geográficos.

Capítulo 4. Exploração de Dados e Análise Espacial

A análise exploratória dos dados espaciais é relevante na medida em que fornece informação valiosa que poderá orientar na escolha dos métodos estatísticos (Waller e Gotway, 2004).

Através da análise espacial é possível transformar dados brutos em informação útil, dada a capacidade de manipular dados espaciais, com o intuito de extrair significado adicional como resultado. Gatrell e Bailey (1995) descrevem três etapas de desenvolvimento que fazem parte da análise espacial, sendo elas: i) visualização, ii) análise exploratória de dados e iii) construção de um modelo. A análise exploratória incide sobretudo na representação gráfica dos dados, porém no caso de informação espacial aplica-se o uso de mapas temáticos de modo a facilitar a interpretação da distribuição geográfica dos dados (Brunsdon, 2008).

Em saúde, no caso do estudo de doenças, o nível de agregação tende a ser maior, não apenas devido às necessidades de confidencialidade dos dados, mas também para aumentar a robustez estatística das taxas (Roquette, *et al.*, 2018). Na presente secção, segue-se a análise e exploração preliminar dos dados, de modo a conceber uma ideia geral da distribuição espacial alusiva a dadores efetivos e doentes transplantados.

4.1. Dadores de órgãos efetivos

É nos hospitais do SNS que dão entrada todos os anos doentes com lesões cerebrais catastróficas com potencial evolução para morte cerebral, i.e., possíveis dadores falecidos de órgãos. Nestes casos, infelizmente é o fim de uma vida, no entanto para alguns doentes é o início de uma nova esperança. Os hospitais pertencentes à Rede Nacional de Coordenação de Colheita e Transplantação, têm um papel fulcral, através da atuação dos coordenadores hospitalares de doação na deteção destes possíveis dadores e envolvimento em todo o processo até à efetivação da doação.

Os dadores têm origem nas várias instituições hospitalares em território nacional, que se articulam com um GCCT de acordo com a sua área geográfica. Existem diversos fatores com impacto na doação de órgãos; no caso da doação após a morte cerebral, um destes fatores é a existência ou não de serviço de neurocirurgia, pelo que na análise efetuada se consideraram dois grupos de hospitais: “com” e “sem” neurocirurgia¹⁷.

No ano de 2016, nos hospitais do SNS com neurocirurgia (Figura 4.1) os Hospitais da Universidade de Coimbra (Centro Hospitalar e Universitário de Coimbra) (51) e o Hospital de São José (Centro Hospitalar Universitário Lisboa Central) (40) destacavam-se pelo elevado número de dadores em morte cerebral, por outro lado no Hospital de Santa Cruz (Centro Hospitalar de Lisboa Ocidental), Hospital de São Francisco Xavier (Centro Hospitalar de Lisboa Ocidental), Hospital Dona Estefânia (Centro Hospitalar Universitário Lisboa Central) e Hospital Curry Cabral (Centro Hospitalar Universitário Lisboa Central) não houve dadores efetivos. Quanto ao ano de 2017, constatou-se um aumento geral do número de dadores efetivos com destaque para o Hospital de São João (Centro Hospitalar Universitário de São João) (93), Hospitais da Universidade de Coimbra (Centro Hospitalar e Universitário de Coimbra) (84), Hospital de São José (Centro Hospitalar Universitário Lisboa Central) (57) e Hospital de Santa Maria (Centro Hospitalar Universitário Lisboa Norte) (46). Ainda no presente ano averigua-se uma ausência de doação no Centro Hospitalar Universitário Cova da Beira, Hospital Curry Cabral (Centro Hospitalar Universitário Lisboa Central), Hospital de Santa Cruz (Centro Hospitalar Lisboa Ocidental) e Hospital de Santa Marta (Centro Hospitalar Universitário de Lisboa Central).

¹⁷ Entende-se como a especialidade médica hospitalar que assegura o diagnóstico e tratamento cirúrgico das doenças do sistema nervoso central, periférico e autónomo.

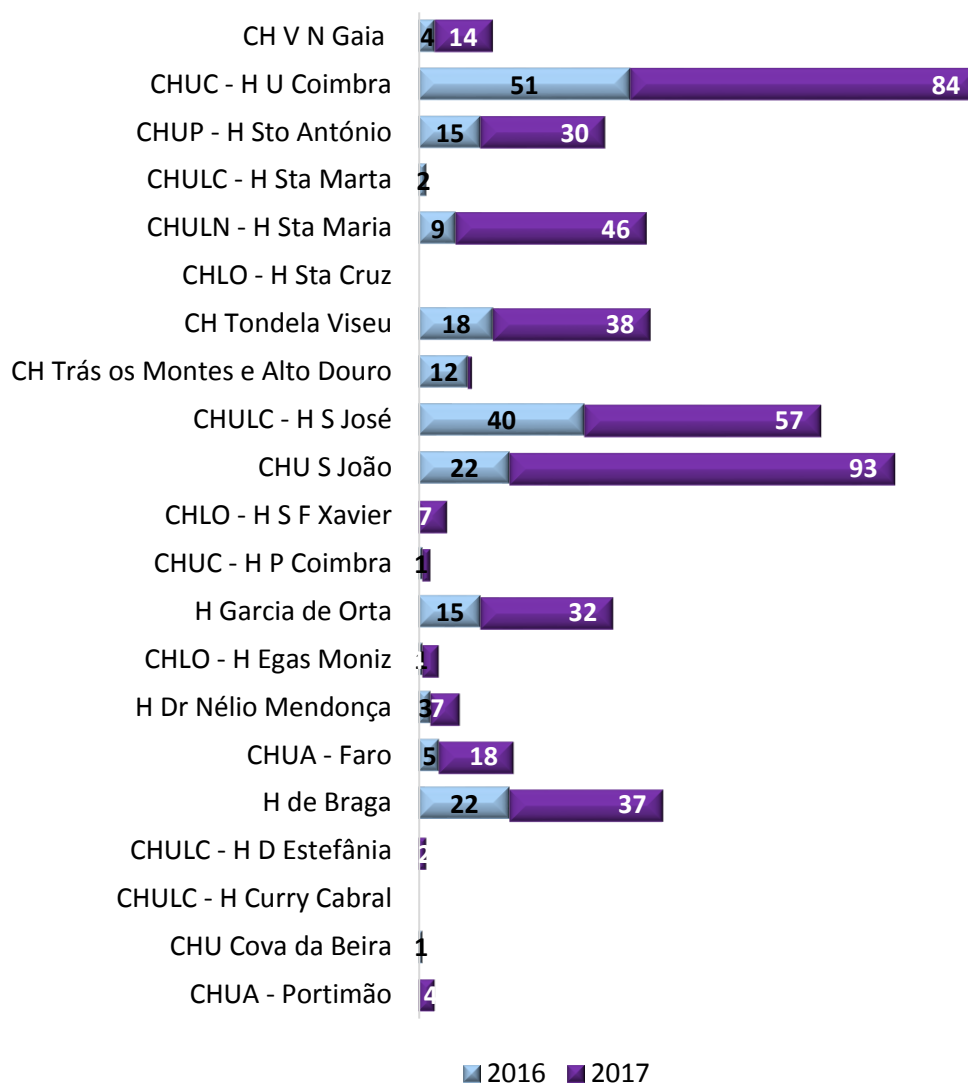


Figura 4.1. Número total de dadores efetivos nos hospitais com neurocirurgia.

Quanto a hospitais sem a especialidade de neurocirurgia (Figura 4.2), de um modo geral também ocorreu um aumento da doação no ano de 2017 face a 2016. Em ambos os anos, o Centro Hospitalar de Entre o Douro e Vouga e o Hospital Professor Doutor Fernando da Fonseca registaram os valores mais elevados de doação.

Porém, no ano de 2016 dá-se uma ausência de doação na Unidade Local de Saúde do Nordeste, Hospital Vila Franca de Xira, Unidade Local de Saúde do Norte Alentejano e Hospital da Horta. No ano de 2017, também é possível constatar ausência de doação no Centro Hospitalar do Barreiro Montijo e no Hospital do Divino Espírito Santo.

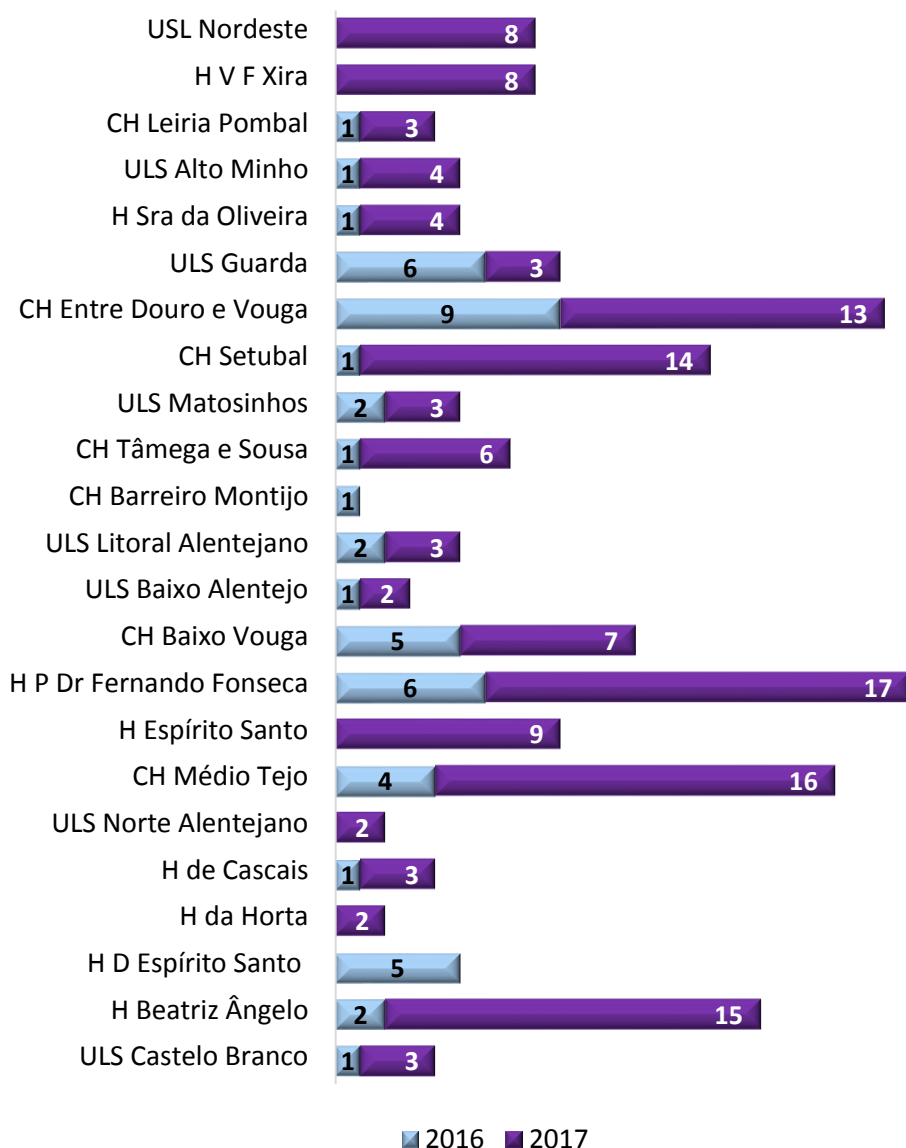


Figura 4.2. Número total de dadores efetivos nos hospitais sem neurocirurgia.

Segundo o Manual de Gestão para a Doação e Transplantação da *DTI Foundation*, a relação preconizada como ideal entre número de dadores efetivos e número de camas de UCI é de 1 para 1, ou seja, 1 dador efetivo por cama de UCI por ano. Transportando este padrão para a realidade portuguesa no ano 2017, constou-se que mais de 70% dos hospitais portugueses com neurocirurgia encontravam-se aquém desta relação (Quadro 4.1).

Quadro 4.1. Relação dos dadores efetivos por camas de UCI, nos hospitais com neurocirurgia.

Dadores/Cama de UCI	Relação	Nº de Hospitais
< 0,09	Baixa	5
0,10 – 0,90	Média Baixa	3
0,91 – 1,48	Média (ideal)	2
1,49 – 3,74	Média Alta	2
≥ 3,75	Alta	1

Geograficamente é possível observar (Figura 4.3) a variabilidade de atividade nos diferentes municípios onde se localizam os hospitais com neurocirurgia. Os municípios que apresentam uma relação média (sendo o preconizado como ideal, 1 dador por cama de UCI ao ano), são o município do Porto através do Centro Hospitalar Universitário de São João, e o município de Almada com o Hospital Garcia de Orta. Numa relação considerada alta, encontra-se no município de Coimbra o Hospital Universitário de Coimbra (Centro Hospitalar e Universitário de Coimbra), e numa relação considerada média alta depara-se o município de Braga com o Hospital de Braga e no município de Viseu o Centro Hospitalar Tondela Viseu. Municípios onde esta relação é considerada média baixa ou baixa, depara-se o Centro Hospitalar de Vila Nova de Gaia Espinho, Centro Hospitalar Universitário do Algarve – Portimão, Centro Hospitalar de Trás os Montes e Alto Douro, Centro Hospitalar Cova da Beira, Hospital de Santa Cruz (Centro Hospitalar de Lisboa Ocidental), Centro Hospitalar Leiria Pombal e Hospital Distrital de Santarém.

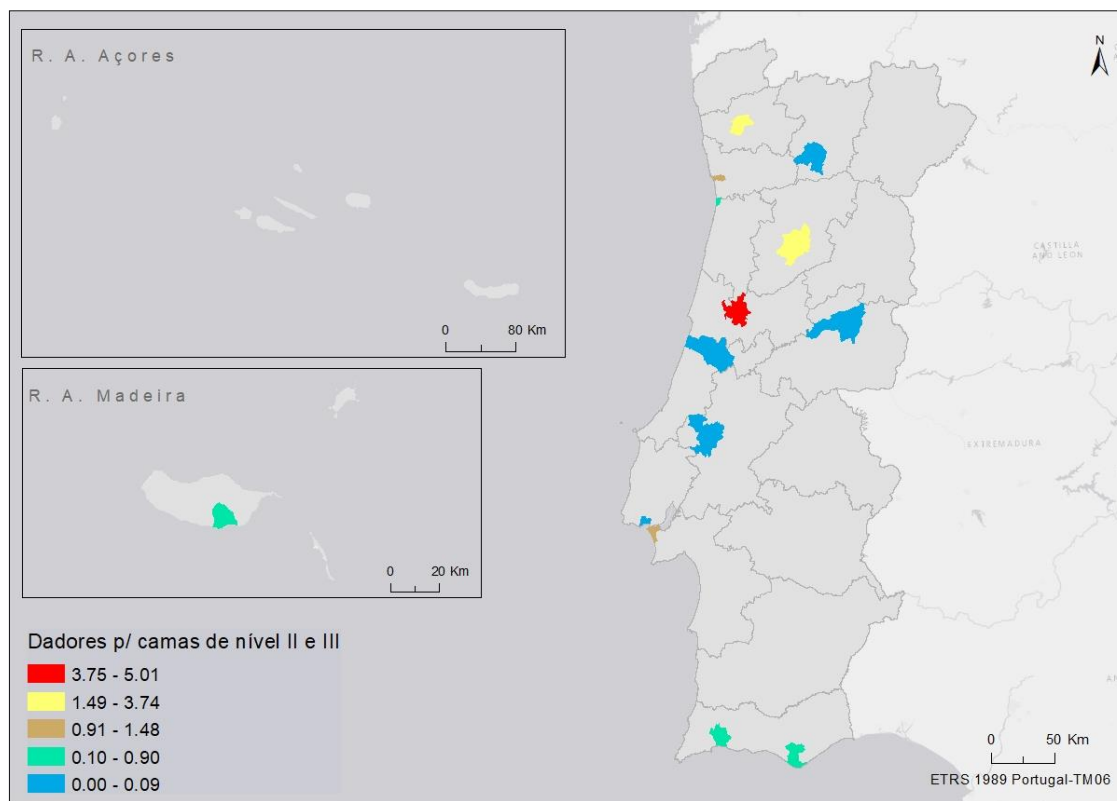


Figura 4.3. Dados efetivos por camas de cuidados intensivos e intermédios, nos hospitais com neurocirurgia ao nível do município

Face aos hospitais do SNS sem a especialidade de neurocirurgia, o padrão da realidade portuguesa do ano 2017, assemelha-se à dos hospitais com neurocirurgia. Constatou-se que a maioria dos hospitais portugueses sem neurocirurgia encontram-se aquém da relação preconizada como ideal (Quadro 4.2).

Quadro 4.2. Relação dos dados efetivos por camas de UCI, nos hospitais sem neurocirurgia

Dadores/Cama de UCI	Relação	Nº de Hospitais
< 0,06	Baixa	5
0,07 – 0,79	Média Baixa	14
0,80 – 1,23	Média (ideal)	2
1,24 – 1,75	Média Alta	0
≥ 1,76	Alta	2

Geograficamente é possível constatar a variabilidade de atividade nos diferentes municípios onde se localizam os Hospitais sem neurocirurgia (Figura 4.4). Apenas dois municípios se encontram próximo do padrão ideal, sendo estes o município de Bragança onde se localiza a Unidade Local de Saúde do Nordeste e o município da Amadora com o Hospital Professor Doutor Fernando da Fonseca. Com uma relação considerada alta, encontrar-se o Centro Hospitalar de Setúbal e o Hospital de Abrantes. Com uma relação considerada baixa, destaque-se o Hospital Distrital de Santarém, o Hospital Divino Espírito Santo em Ponta Delgada, o Hospital Distrital de Pombal, Hospital Nossa Sra. Do Rosário e o Hospital de Santo André.

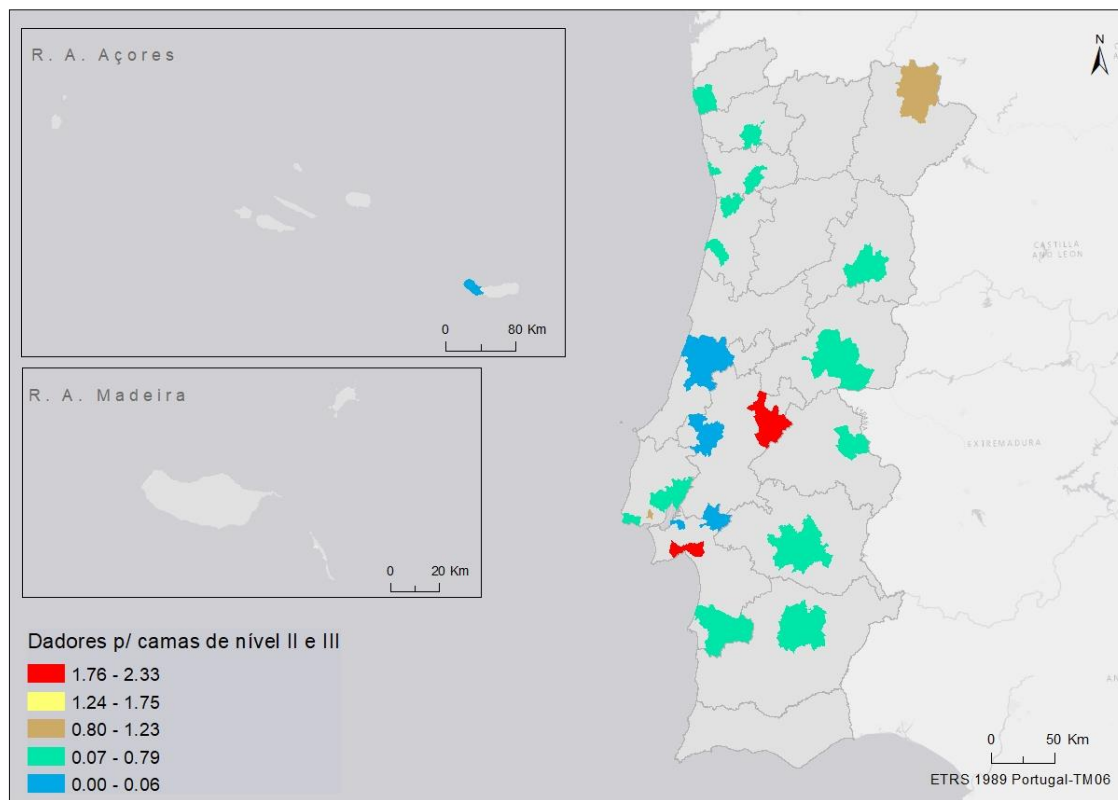


Figura 4.4. Dadores efetivos por camas de cuidados intensivos e intermédios, nos hospitais sem neurocirurgia ao nível do município

Quanto à distribuição dos dadores por município dado o seu local de residência no ano de 2016 (Figura 4.5), um pouco como espectável, facilmente se destacam as duas grandes metrópoles. O maior aglomerado de dadores, sobressai sobretudo na área

metropolitana do Porto. Também a região Centro, Alentejo e Algarve contribuem positivamente para a doação.

Porto é o município com o maior número de doadores (20), seguindo-se Lisboa (15), Valongo (13), Maia (11), Vila Nova de Gaia (10), Gondomar (8), Sintra (8), em 16º lugar Albufeira (3), Braga (3), Guimarães (3), entre outros municípios.

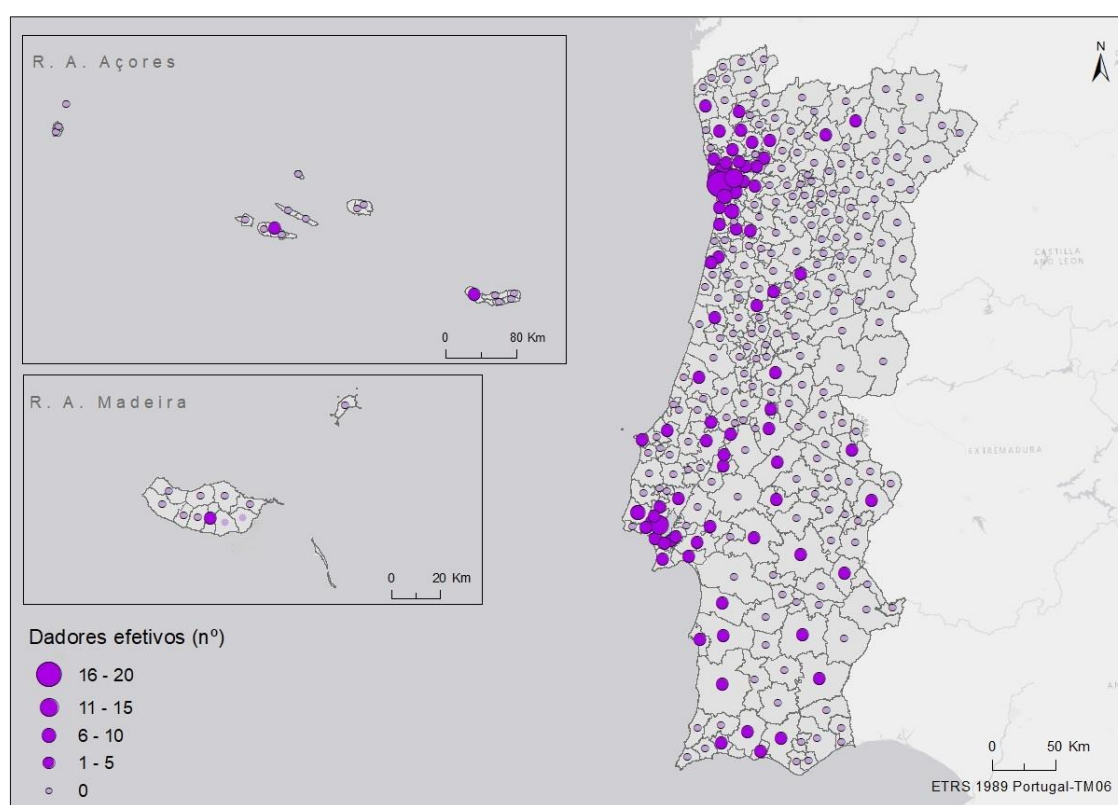


Figura 4.5. Número de doadores efetivos por município no ano de 2016

Relativamente ao ano de 2017 (Figura 4.6), tendo em consideração o aumento do número de doadores, verifica-se uma maior expressividade no território em relação a estes. Este crescimento estende-se por todo o país, com particular impacto nos municípios da região litoral norte.

Os valores mais elevados de doação encontram-se, por ordem decrescente, nos municípios de Sintra (17), Gondomar (12), Lisboa (10), Porto (10), Vila Nova de Gaia (10), Maia (9), Barcelos (5), Coimbra (5), Viseu (5), entre outros municípios.

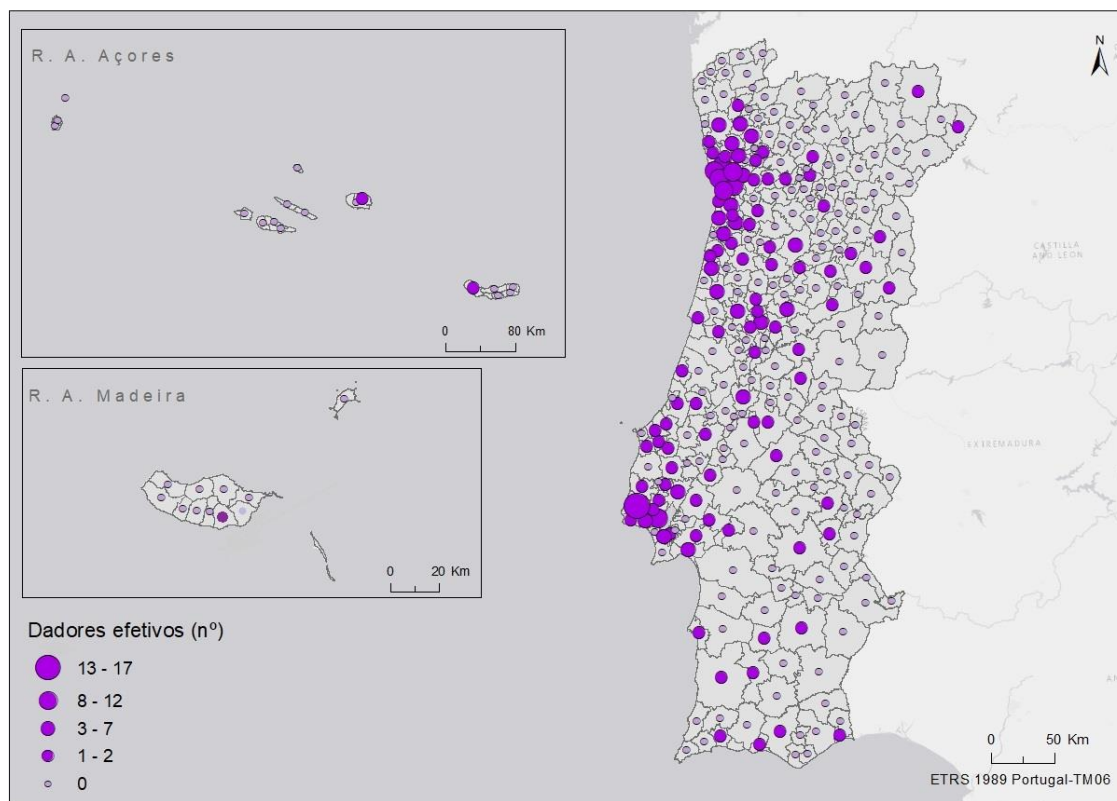


Figura 4.6. Número de doadores efetivos por município no ano de 2017.

A Figura 4.7, numa análise por NUTS II, retrata a taxa de utilização de órgãos face ao número de colheitas nos hospitais pertencente às várias regiões. O Norte apresenta o maior número de órgãos colhidos (384), com uma taxa de utilização de 84 %, seguindo-se a Área Metropolitana de Lisboa (287) e a região Centro (286), com taxas de utilização de 93% a 94%, respetivamente. Com menor significância segue-se o Algarve (30), a região do Alentejo (21), a região autónoma dos Açores (4) e a região autónoma da Madeira (10), com taxas de utilização de 90%, 90%, 100% e 80%, respetivamente.

Quanto à taxa de reprovação¹⁸ (Figura 4.8), a região Norte embora seja a que mais órgãos colhe por ano, também é a segunda região com uma maior taxa de reprovação de órgãos, cerca de 16%. Quanto à Área Metropolitana de Lisboa e região Centro, embora estas apresentem um número total de órgãos colhidos inferior à região Norte, porém, detêm uma taxa de reprovação de apenas 6% a 7% em relação às restantes regiões, o que se reverte em um maior aproveitamento.

¹⁸ Entenda-se como a taxa de não utilização de órgãos.

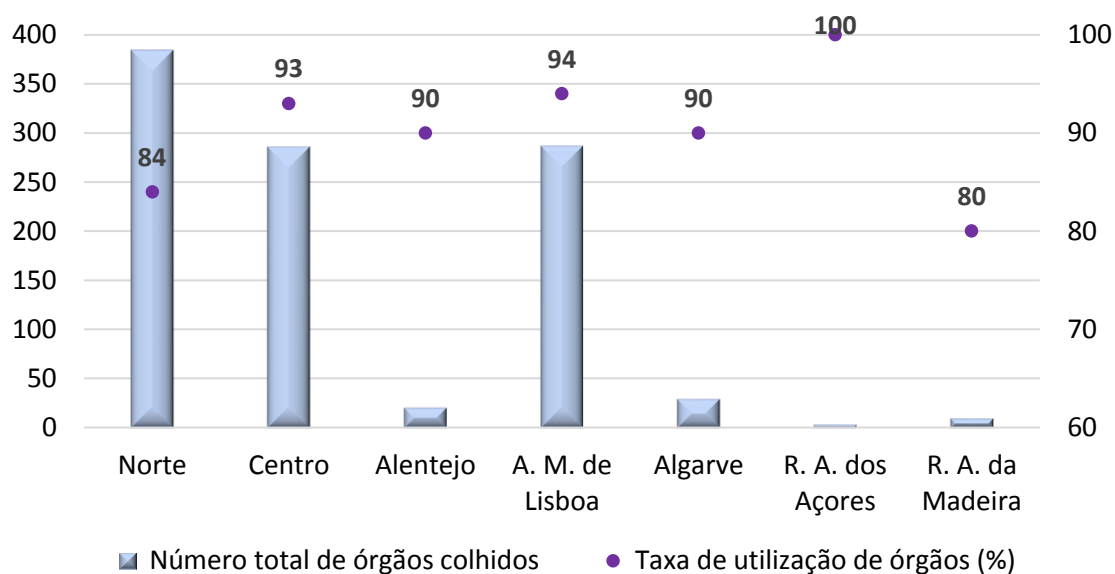


Figura 4.7. Número de órgãos colhidos e respetiva taxa de utilização nos hospitais pertencentes à NUTS II no ano de 2017.

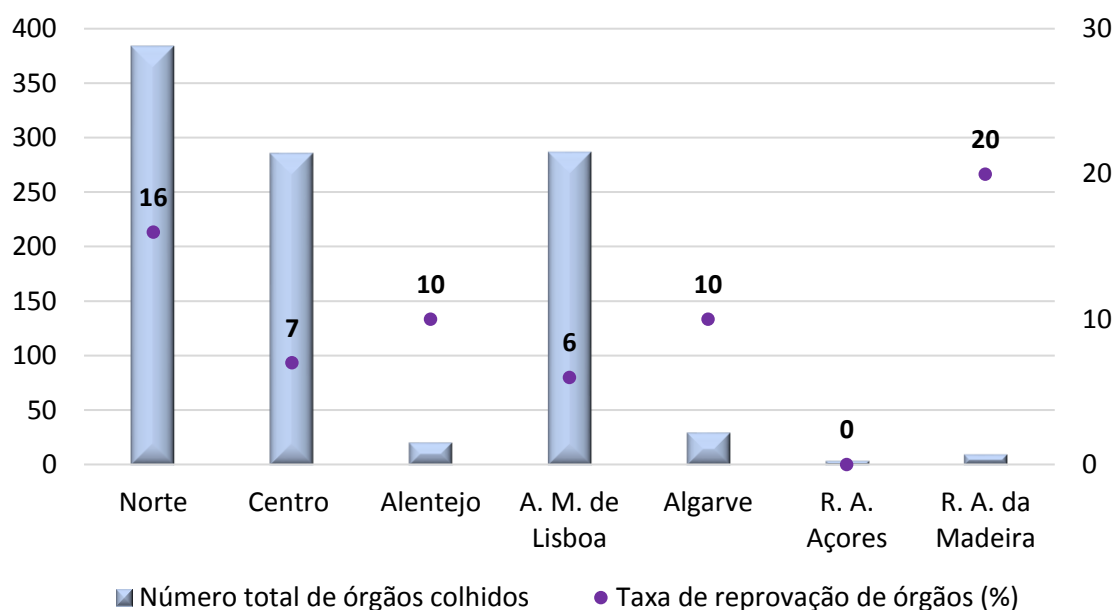


Figura 4.8. Número de órgãos colhidos e respetiva taxa de reprovação de órgãos nos hospitais pertencentes às NUTS II no ano de 2017.

De modo a analisar os municípios de Portugal relativamente a dadores de órgãos efetivos, é relevante iniciar com uma caracterização da população em termos de idade

(Figura 4.9, uma vez que a maioria das pessoas que são dadoras em Portugal possuem idades compreendidas entre os 45 e os 65 anos de idade.

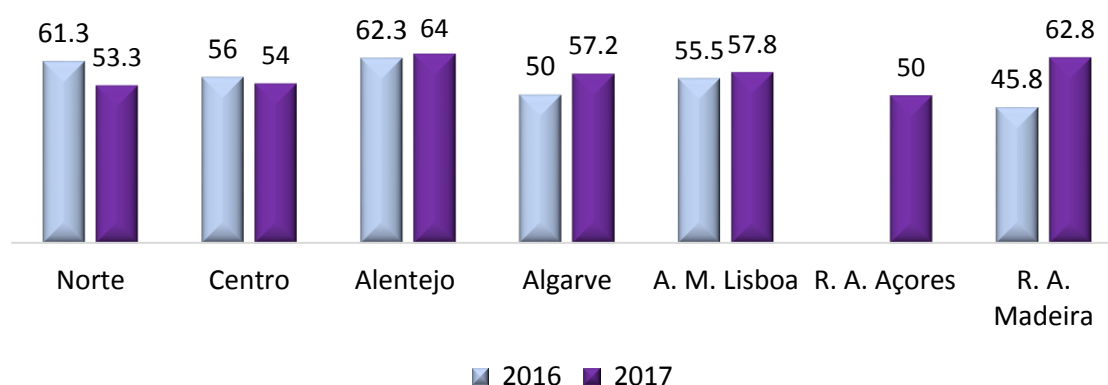


Figura 4.9. Idade média dos dadores efetivos por NUTS II.

A Figura 4.10 representa a população residente em Portugal ao nível do município, com idades compreendidas entre os 45 e os 65 anos de idade. Esta, ainda que se localize maioritariamente nas Áreas Metropolitanas de Lisboa e do Porto, tem maior concentração na região litoral. Fora das metrópoles, destacam-se cidades como Braga, Guimarães e Coimbra. Já no centro do país destaca-se Viseu, Covilhã, Castelo Branco e Évora. Quanto às regiões Autónomas da Madeira e dos Açores destacam-se os municípios do Funchal e de Ponta Delgada, respetivamente.

A Figura 4.11 e a Figura 4.12 representam a análise de uma outra característica dos dadores efetivos, sendo esta o género para os anos 2016 e 2017.

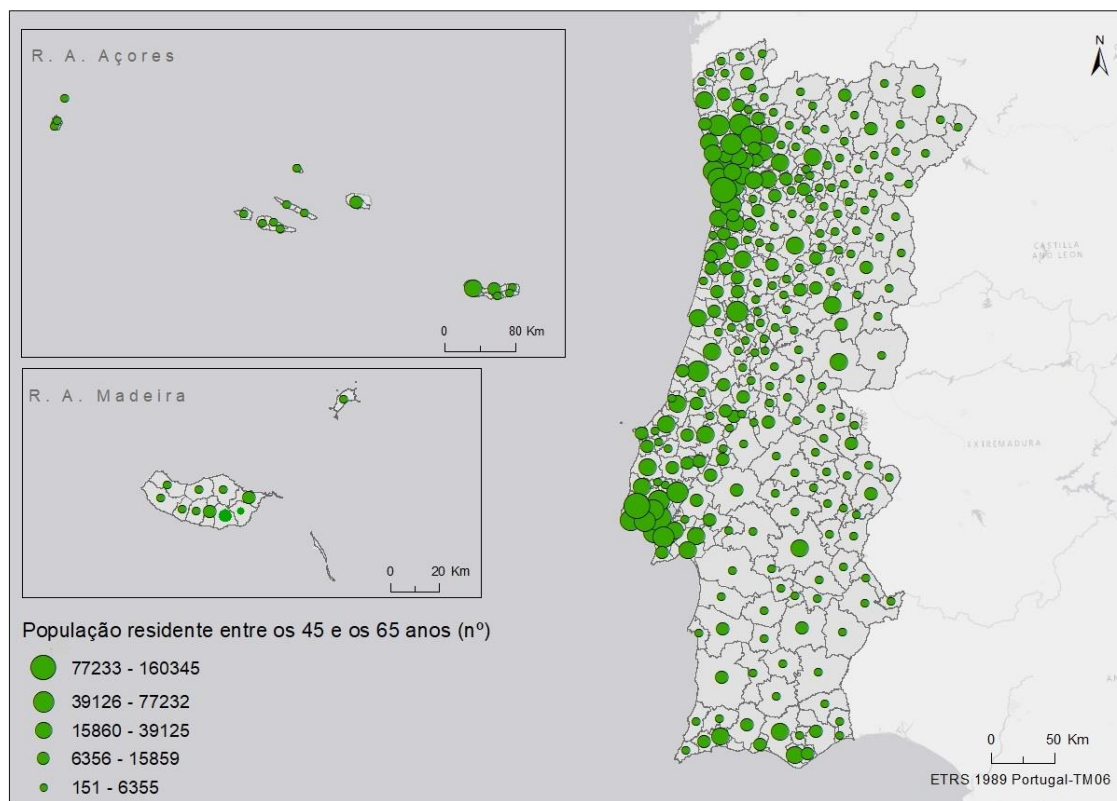


Figura 4.10. População residente entre os 45 e os 65 anos de idade por município.

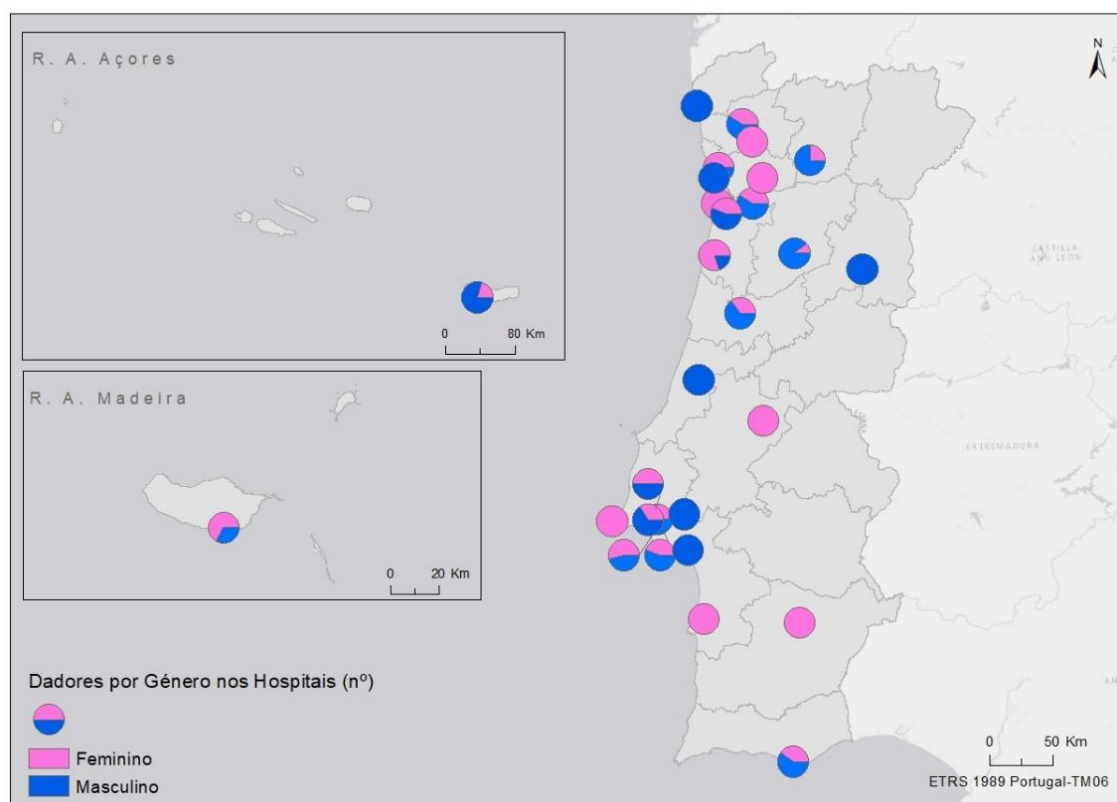


Figura 4.11. Dadores efetivos por género nos hospitais do SNS no ano de 2016.

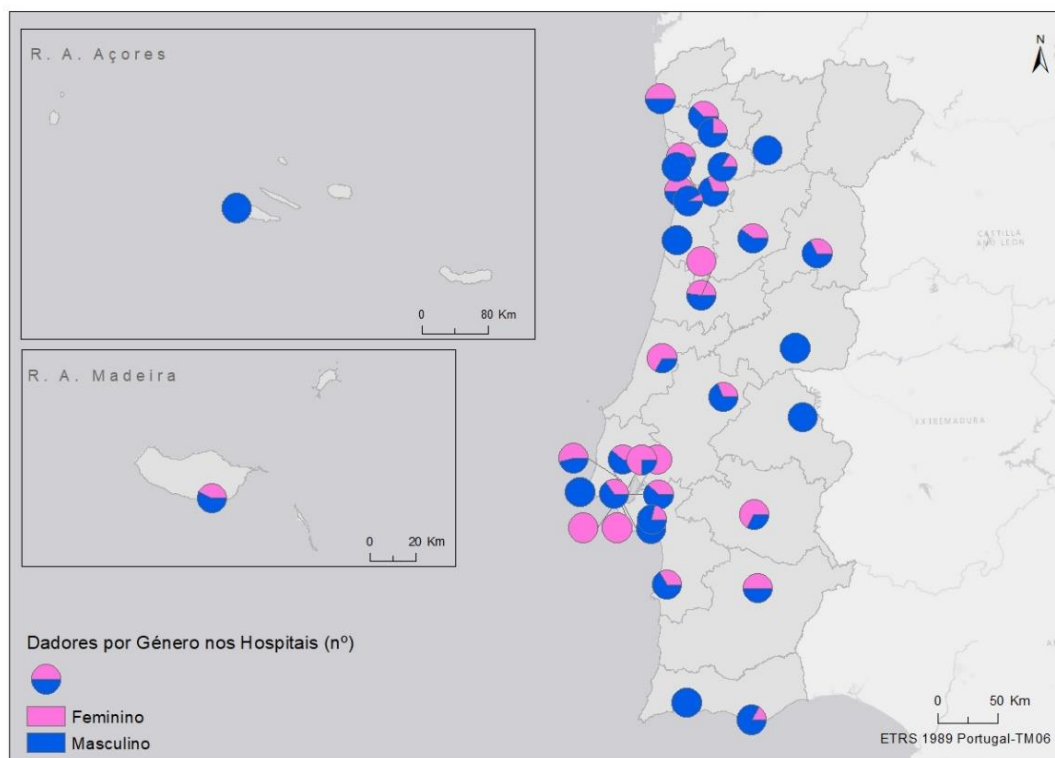


Figura 4.12. Dadores efetivos por género nos hospitais do SNS no ano de 2017.

Face ao número de dadores nos hospitais do SNS, especialmente não é possível constatar nenhum padrão associado ao género, uma vez que este difere ao longo dos anos analisados. No entanto, grosso modo, averigua-se que a maioria dos dadores são do género masculino (Figura 4.13).

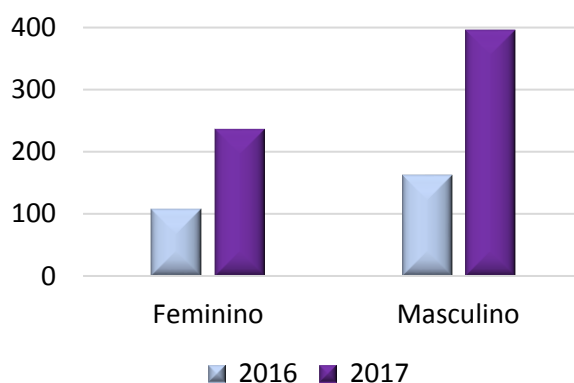


Figura 4.13. Dadores efetivos por género.

Atualmente, com a redução da sinistralidade rodoviária, verifica-se uma alteração do perfil do dador com diminuição do número de dadores com causa de morte de origem traumática. O gráfico da Figura 4.14, exemplifica as causas de morte dos dadores efetivos dos anos em estudo. Os acidentes vasculares cerebrais, especialmente os hemorrágicos, revelam-se a causa de morte que mais contribui para a doação. Ainda dentro das causas médicas, destacam-se os óbitos por anoxia¹⁹, e nas causas traumáticas os traumatismos crânio encefálicos²⁰ de origem não especificada.

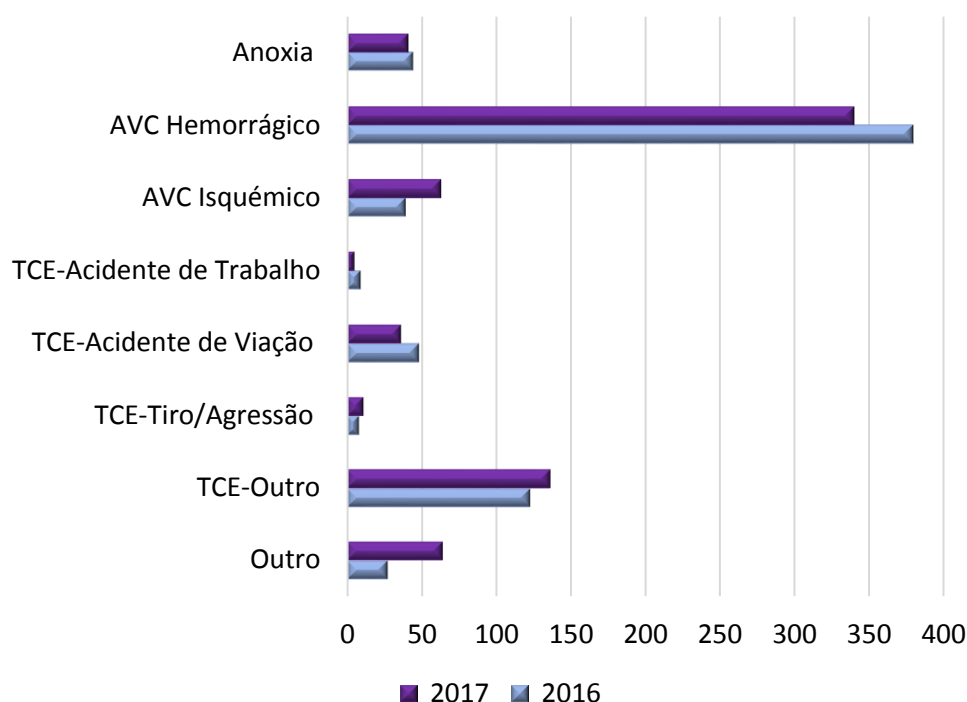


Figura 4.14. Distribuição dos dadores efetivos de acordo com as causas de morte.

Segundo o Instituto Nacional de Estatística²¹ “as doenças do aparelho circulatório continuam a ser a principal causa de morte em Portugal. No conjunto das doenças do aparelho circulatório, evidenciam-se os AVC que representaram 10,8% do total de mortes no país e a doença isquémica do coração com 6,7%.”

¹⁹ Entenda-se a falta de oxigénio no organismo por respiração deficiente.

²⁰ Também conhecido como lesão intracraniana, ocorre quando uma força externa causa um ferimento traumático no cérebro.

²¹ INE, “Causas de Morte – 2015”, edição 2017.

Site: http://www.ine.pt/portal/page/portal/PORTAL_INE/Publicacoes

Segue-se, na Figura 4.15, a proporção de óbitos por doença do aparelho circulatório, nos municípios de Portugal. Estes são uma constante que se faz sentir um pouco por todo o território, em proporções que vão desde os 14,3% aos 48,1% nos vários municípios do país, com particularidade no interior.

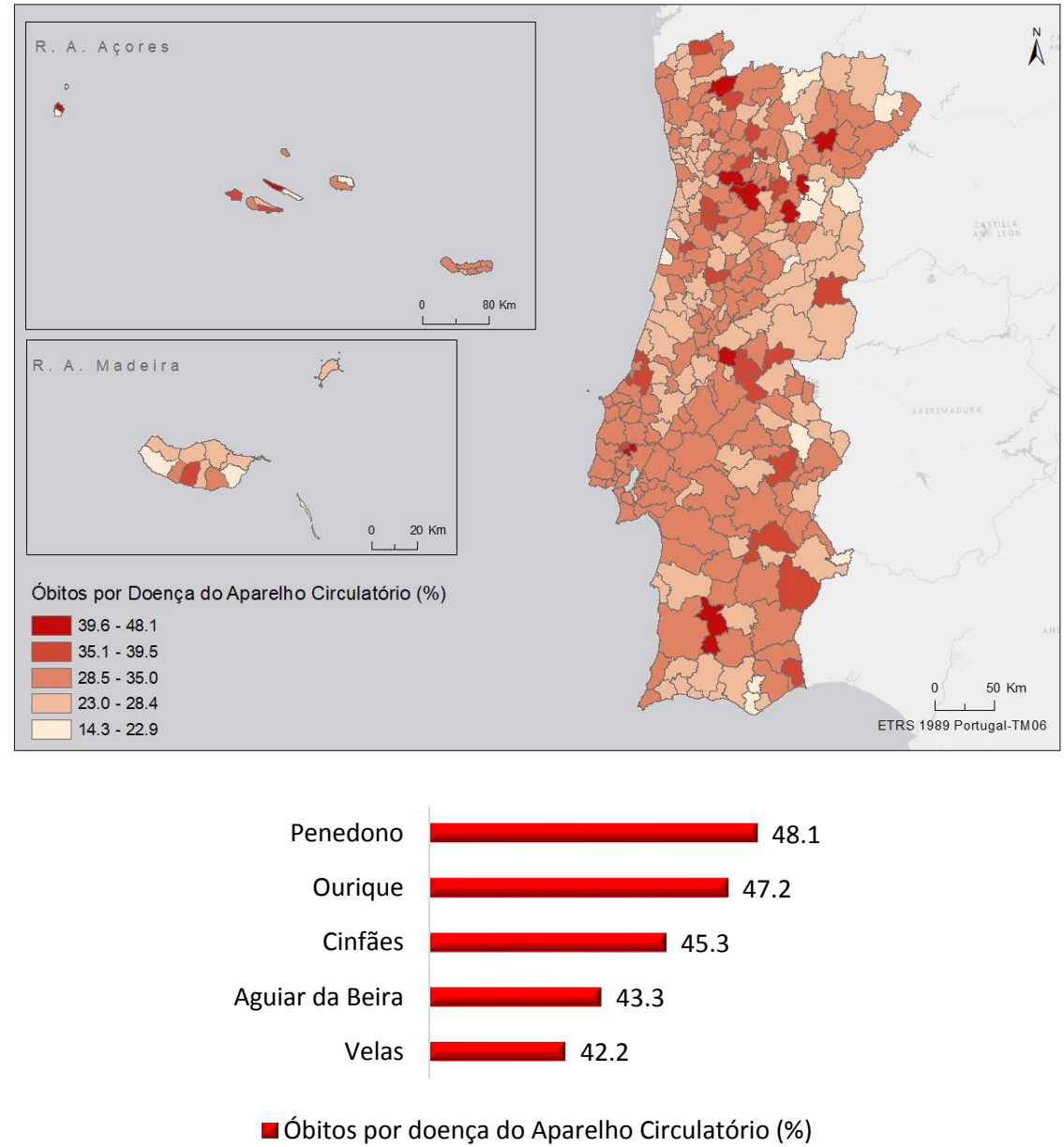


Figura 4.15. Proporção de óbitos por doença do aparelho circulatório nos municípios de Portugal.

Embora numa percentagem mais reduzida (19%) em 2017, também os óbitos por acidente de viação (Figura 4.16) obtiveram significância no que concerne às causas de morte dos dadores, sobretudo em municípios como Loulé, Sobral do Monte Agraço e Vila Nova de Foz Côa.

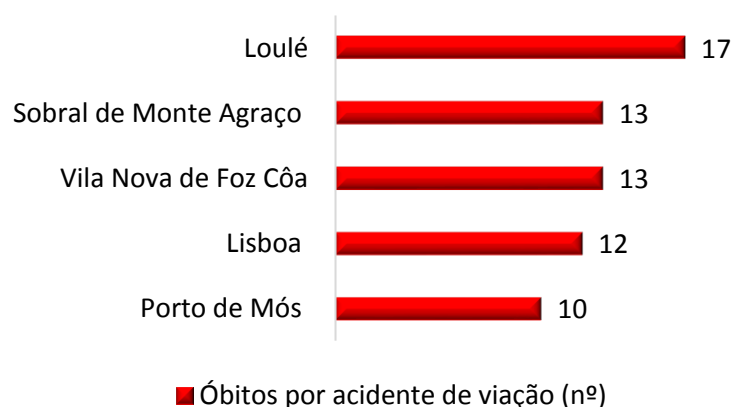
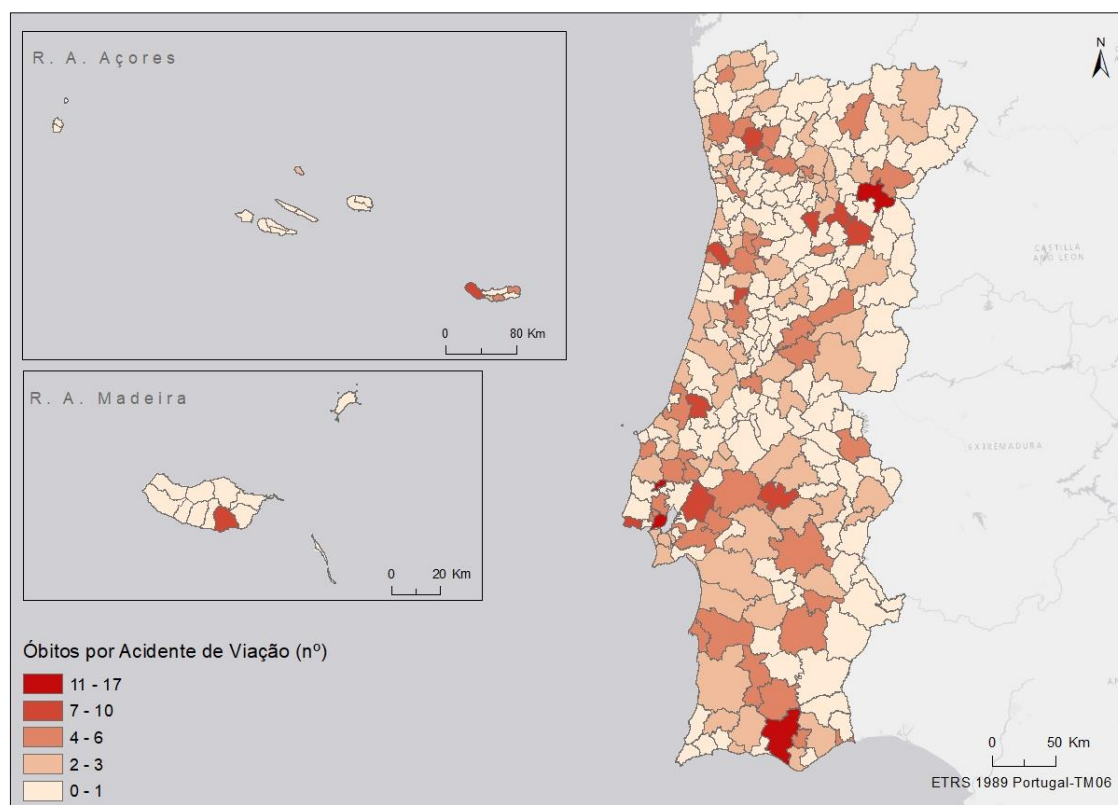


Figura 4.16. Número de óbitos por acidente de viação nos municípios de Portugal.

De modo a perspetivar o fluxo de dadores efetivos, considerou-se a morada dos dadores e a ligação desta ao respetivo hospital onde foi efetivada a doação. Especialmente é possível constatar algumas assimetrias (Figura 4.17). Na maioria dos casos (13 a 17

dadores), o local de residência do dador efetivo encontra-se a uma distância próxima do hospital onde se efetivou a doação, ocorrência esta, sobretudo nos grandes centros urbanos de Lisboa e do Porto. Em casos mais esporádicos (1 dador) verifica-se uma elevada distância da morada do dador face ao respetivo hospital. No caso das ilhas, existem dadores que são residentes nas regiões autónomas, porém, findaram por ser dadores no continente e vice-versa, i.e., é o caso de um dador que tinha por morada o município do Funchal, no entanto acabou por dar entrada como dador no Hospital Universitário de Coimbra.

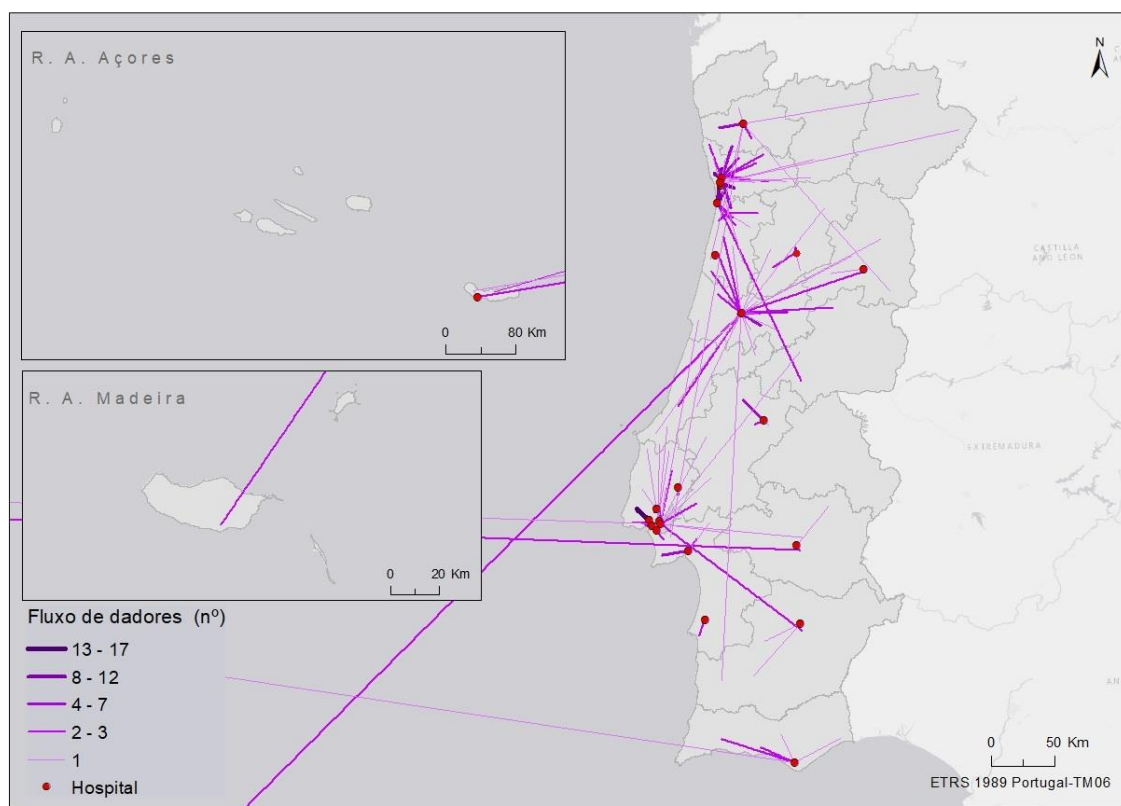


Figura 4.17. Fluxo de dadores face à morada do dador efetivo e respetivo hospital de colheita.

4.2. Doentes transplantados

O Quadro 4.3 é representativo do número de doentes transplantados por hospital no ano de 2016, assim como o respetivo órgão transplantado. Relativamente aos órgãos, o rim é o órgão mais transplantado em Portugal, seguindo-se o fígado, o coração, o

pulmão e o pâncreas. Quanto às unidades de transplantação, o Hospital Universitário de Coimbra apresenta-se como pioneiro na transplantação de órgãos no ano de 2016, seguindo-se o Hospital de Santo António no Porto e o Hospital Curry Cabral em Lisboa.

Quadro 4.3. Número de doentes transplantados por hospital e respetivo órgão no ano de 2016.

Hospital / Órgão	Rim	Pâncreas	Fígado	Pulmão	Coração	Total
H Santa Cruz	57				5	62
H Garcia de Orta	12		1			13
H São João	49				4	53
H Santo António	88	13	41			142
H Curry Cabral	36	6	58			100
H Santa Maria	25					25
H Santa Marta	1			30	5	36
H U. de Coimbra	105		65		20	190
Total	373	19	165	30	34	

Relativamente ao ano de 2017 (Quadro 4.4), o rim mantém-se como o órgão mais transplantado em Portugal, seguindo-se novamente o fígado, o pulmão, o coração e o pâncreas. O Hospital Universitário de Coimbra, Hospital Curry Cabral e o Hospital de Santo António, mantêm-se como as principais unidades de transplantação dado o elevado número de cirurgias deste cariz, realizadas anualmente.

Quanto ao local de residência dos doentes transplantados (Figura 4.18) sobressaem, como seria espectável, os grandes centros urbanos da Área Metropolitana do Porto e da Área Metropolitana de Lisboa, seguindo-se cidades como Coimbra, o que se dilucida pelo facto de serem os municípios com mais população.

Quadro 4.4. Número de doentes transplantados por hospital e respetivo órgão no ano de 2017.

Hospital / Órgão	Rim	Pâncreas	Fígado	Pulmão	Coração	Total
H Santa Cruz	59				5	64
H Garcia de Orta	15					15
H São João	76				5	81
H Santo António	103	13	56			172
H Curry Cabral	65	13	127			205
H Santa Maria	74					74
H Santa Marta				60	8	68
H U. de Coimbra	140		78		28	246
Total	532	26	261	60	46	

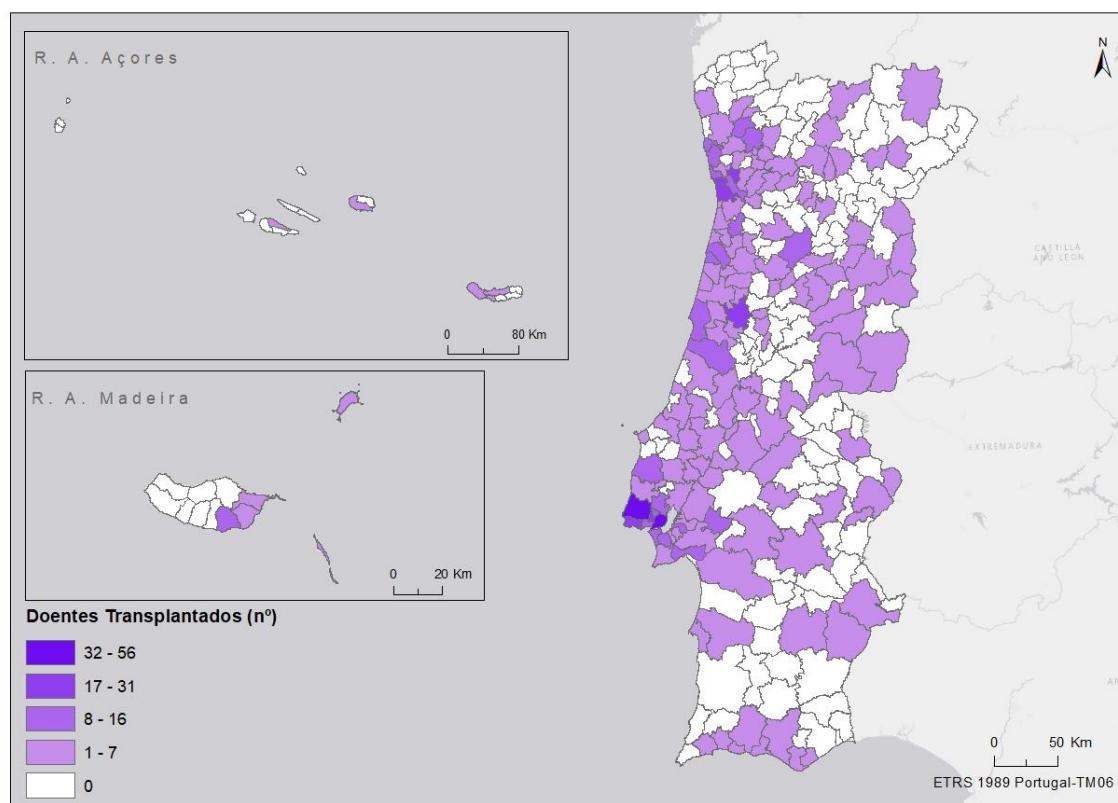


Figura 4.18. Número de doentes transplantados por município no ano de 2017.

Quanto à incidência²² de doentes transplantados em 2016, considerando a população residente por município, o panorama altera-se uma vez que a incidência passa a ser elevada sobretudo no interior do continente (Figura 4.19). O mesmo se sucede relativamente ao ano de 2017 (Figura 4.20), com destaque novamente para o interior, e de configuração mais acentuada. Tal facto leva-nos a constatar uma resposta positiva às necessidades dos doentes transplantados residentes um pouco por todo o país, em especial no interior.

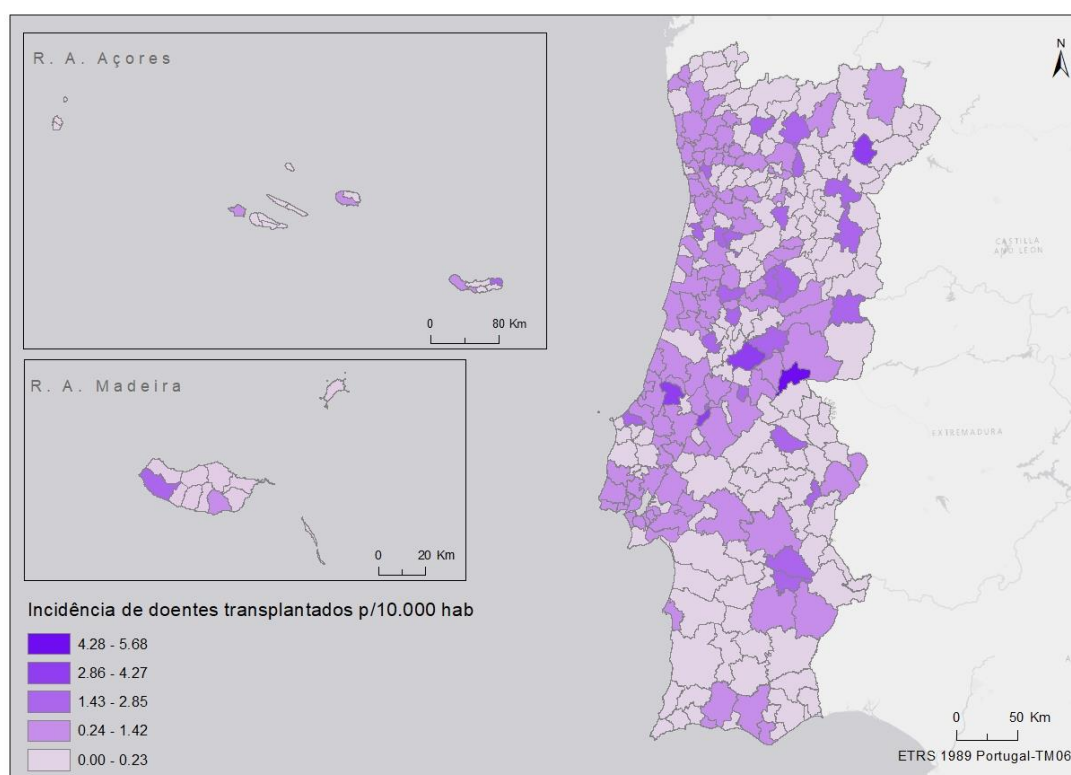


Figura 4.19. Incidência de doentes transplantados face à sua origem nos municípios de Portugal, no ano de 2016.

²² Entenda-se como o:

$$\frac{\text{número de doentes transplantados por município}}{\text{número de população residente por município}} \times 10.000 \text{ habitantes}$$

“refere-se à ocorrência de novos casos dentro de um período de tempo especificado” (Waller e Gotway, 2004).

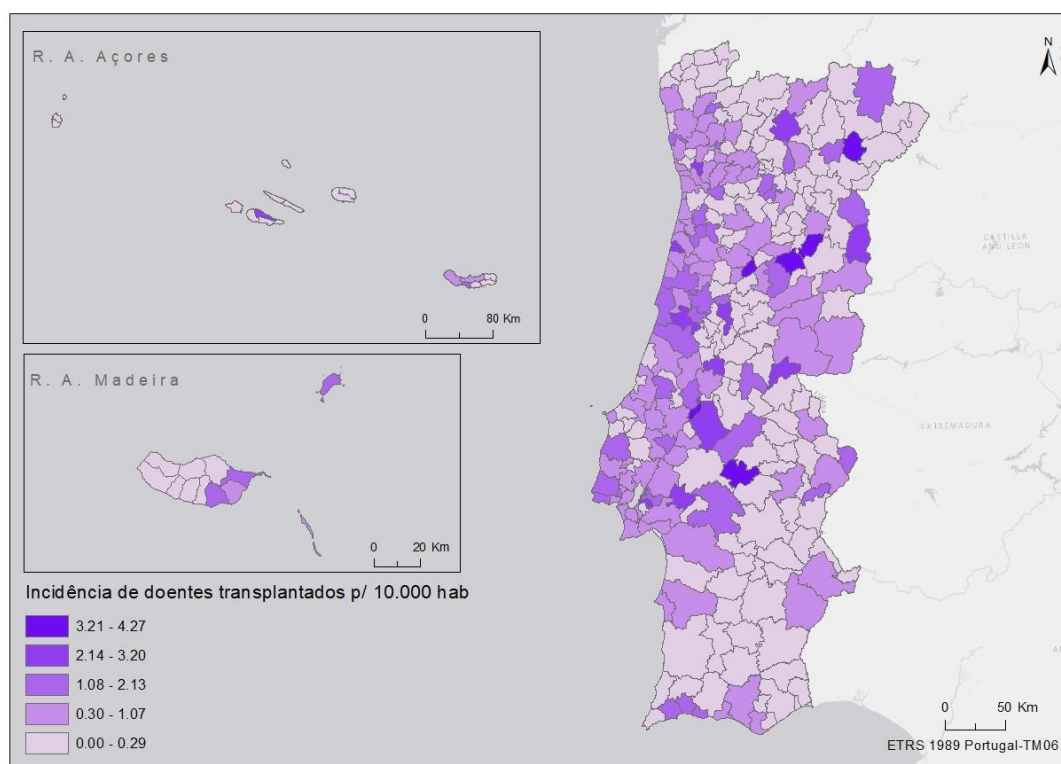


Figura 4.20. Incidência de doentes transplantados face à sua origem nos municípios de Portugal, no ano de 2017.

Em Portugal o rim é o órgão mais transplantado, assim como aquele que detém um maior número de unidades de transplantação. Somente a doentes que aguardam por um transplante renal, é-lhes dada a hipótese de optarem pela inscrição até dois hospitais que detenham unidade de transplantação renal.

A Figura 4.21 demonstra o fluxo do acesso de doentes transplantados renais em 2017, com base no seu local de residência e a respetiva deslocação à unidade de transplantação, onde realizaram a cirurgia.

Relativamente às unidades de transplantação do Porto, é possível constatar que a maioria dos doentes residem próximos desta, onde o mesmo se sucede nas unidades de transplantação de Lisboa. Quanto à unidade de transplantação do Centro Hospitalar e Universitário de Coimbra, apura-se uma maior dispersão geográfica dos doentes, dado o acesso destes ao centro de transplante. Esta é uma unidade que transplanta doentes originários um pouco de todo o país.

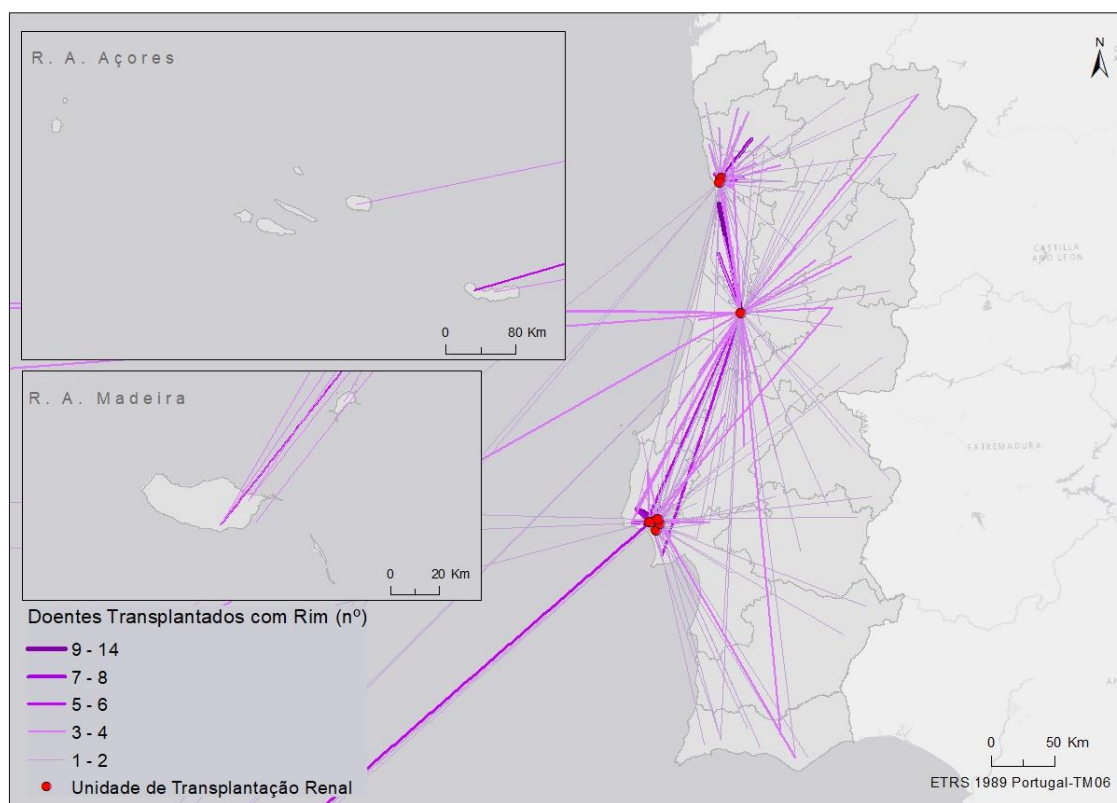


Figura 4.21. Fluxos de acesso de doentes transplantados com rim aos hospitais com unidade de transplantação Renal

No que concerne ao fígado (Figura 4.22), o maior número de fluxos concentra-se sobretudo na faixa litoral do continente, onde permanece na grande maioria dos casos a proximidade às unidades de transplantação. Na região Norte, é o caso do distrito de Braga, onde o fluxo de acesso de doentes à unidade de transplantação hepática do Centro Hospitalar Universitário do Porto é elevado (11 a 13 doentes). Quanto à região Centro, é no distrito de Coimbra que se encontra localizada uma das três unidades de transplantação hepática, onde o fluxo de doentes (11 a 13 doentes) é proveniente grosso modo do distrito de Aveiro e do próprio distrito de Coimbra. No distrito de Lisboa, os doentes transplantados na unidade hepática do Centro Hospitalar Universitário de Lisboa Central, têm residência na grande maioria o distrito de Lisboa (14 a 16 doentes), onde permanece a proximidade ao centro de transplante. Em casos mais esporádicos (1 a 4 doentes), onde verifica-se um maior distanciamento do local de residência face à unidade de transplantação.

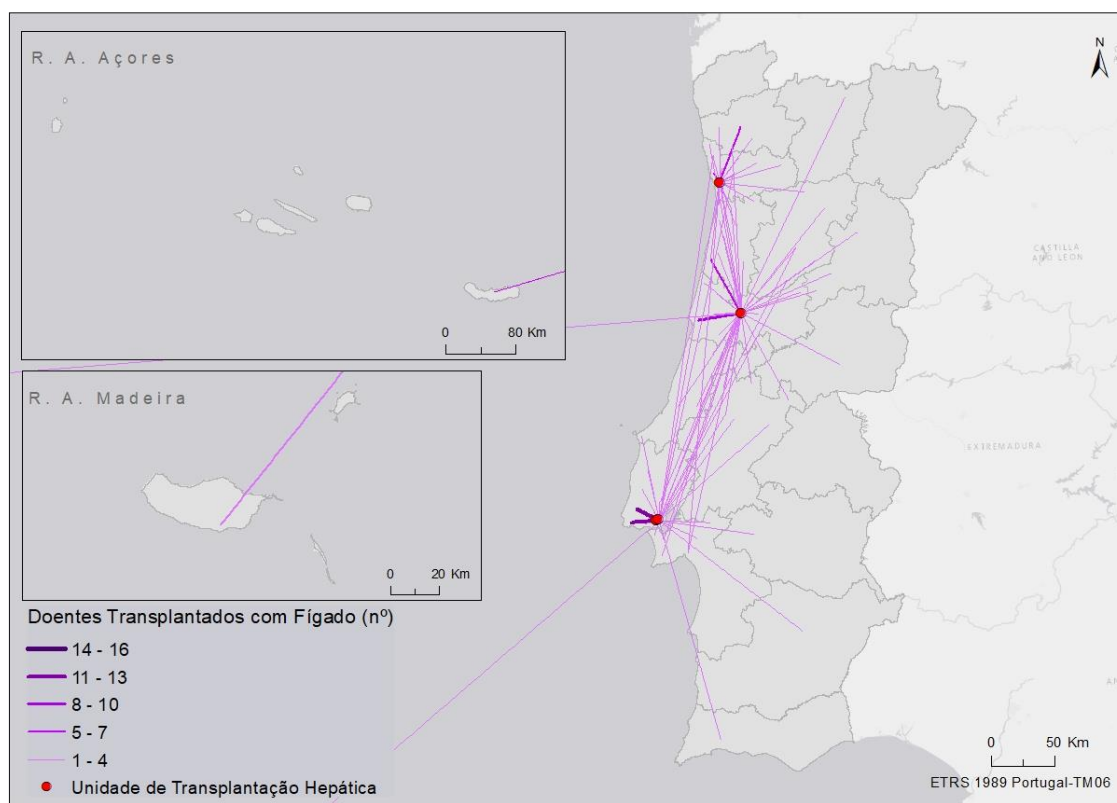


Figura 4.22. Fluxos de acesso de doentes transplantados com fígado aos hospitais com Unidade de Transplantação Hepática

Relativamente a doentes transplantados de coração (Figura 4.23), o fluxo de acesso é de um modo geral, disperso. É particularmente na unidade de transplantação do Centro Hospitalar e Universitário de Coimbra, onde o acesso de doentes representa uma maior dispersão geográfica face à acessibilidade ao centro de transplante. Na região do Alentejo e Algarve verifica-se uma ausência de doentes transplantados por coração.

Quanto a doentes transplantados de pulmão (Figura 4.24), o fluxo de acesso é denso no que respeita a doentes residentes na região Norte e Centro do continente, o que é explicado por existir apenas um centro de transplantação pulmonar, no Centro Hospitalar Universitário Lisboa Central (Hospital de Santa Marta).

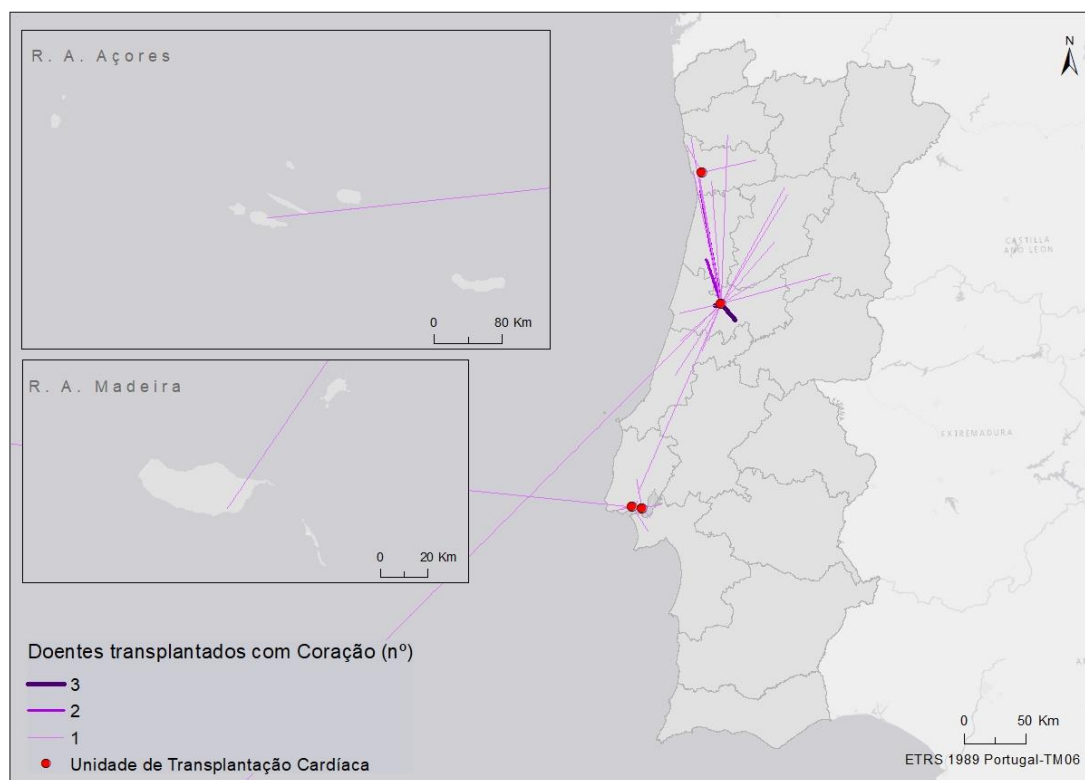


Figura 4.23. Fluxos do acesso de doentes transplantados com coração aos hospitais com Unidade de Transplantação cardíaca

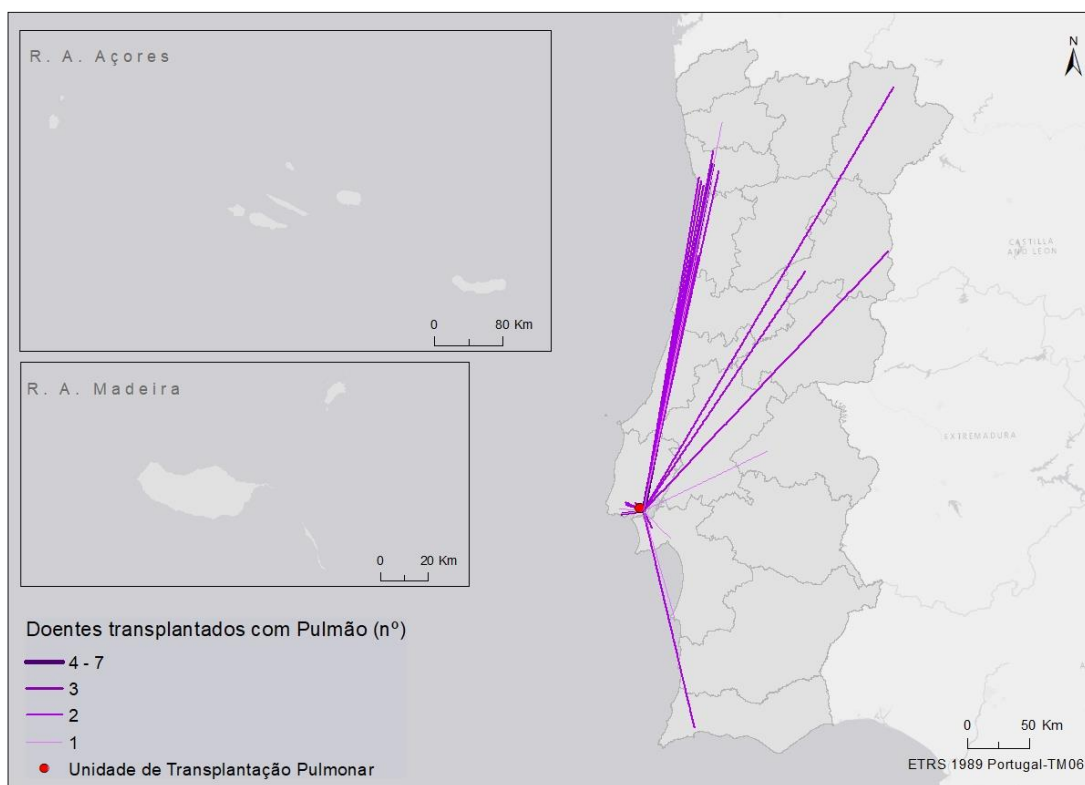


Figura 4.24. Fluxos do acesso de doentes transplantados com pulmão ao hospital com Unidade de Transplantação Pulmonar

Em comparação com as restantes tipologias de transplantação, o número de doentes transplantados de pâncreas anualmente (Figura 4.25) é o mais reduzido, o que se reverte numa ocorrência de fluxos como menor expressão. Face às unidades de transplantação pancreática, é de constatar que os doentes optam pela unidade mais próxima ao seu local de residência para realizar a cirurgia.

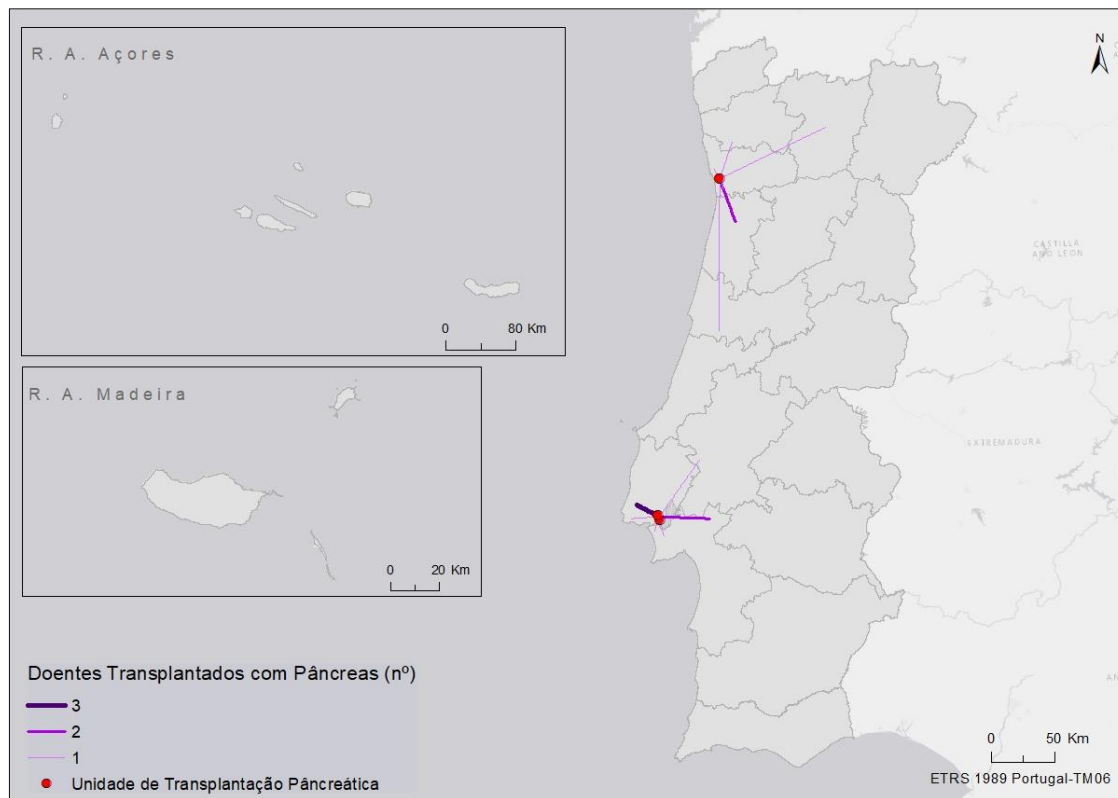


Figura 4.25. Fluxos do acesso de doentes transplantados com pâncreas aos hospitais com Unidade de Transplantação pancreática

Capítulo 5. Análise dos Resultados

Ao analisar padrões espaciais de doença, é também possível estudar padrões espaciais associado à exposição de potenciais. Na verdade, na maioria dos casos é necessário considerar as informações de exposição espacial de modo a entender a variação espacial observada (Waller e Gotway, 2004).

No presente capítulo, são discutidos os resultados obtidos pela aplicação dos métodos estatísticos, apresentados anteriormente no capítulo 4. Para um melhor entendimento, os resultados encontram-se repartidos em dois subcapítulos: um primeiro referente à doação, e um segundo acerca da transplantação.

5.1. Dadores efetivos

5.1.1. Índice Local de Moran

Com o intuito de perspetivar quais as diferenças entre *clusters* face à magnitude dos dadores, seguem-se os valores do Índice de Moran.

Através do nível espacial de agregação dos dados, é possível constatar disparidades na distribuição geográfica dos dadores efetivos (Figura 5.1). Um pouco por todo o continente, são os *clusters* de magnitude reduzida os que mais se destacam no território. As classes mais representadas possuem valores de índice negativo (-0,09 e -0,44), o que indica uma correlação indireta, portanto uma concentração espacial menor de dadores face à população residente nestes municípios. Nesses locais é possível destacar municípios como Cascais, Mafra, Montijo, Tomar, Coimbra, Vila Verde e Vila do Conde.

Quanto a valores de índice positivo, ou seja, uma correlação direta de valores mais elevados, estes concentram-se em municípios como Gondomar, Maia, Matosinhos, Valongo, Porto e Vila Nova de Gaia. Tal facto, leva a afirmar que os *clusters* de maior

magnitude face a dadores estão localizados na Área Metropolitana do Porto, seguindo-se o município de Sintra.

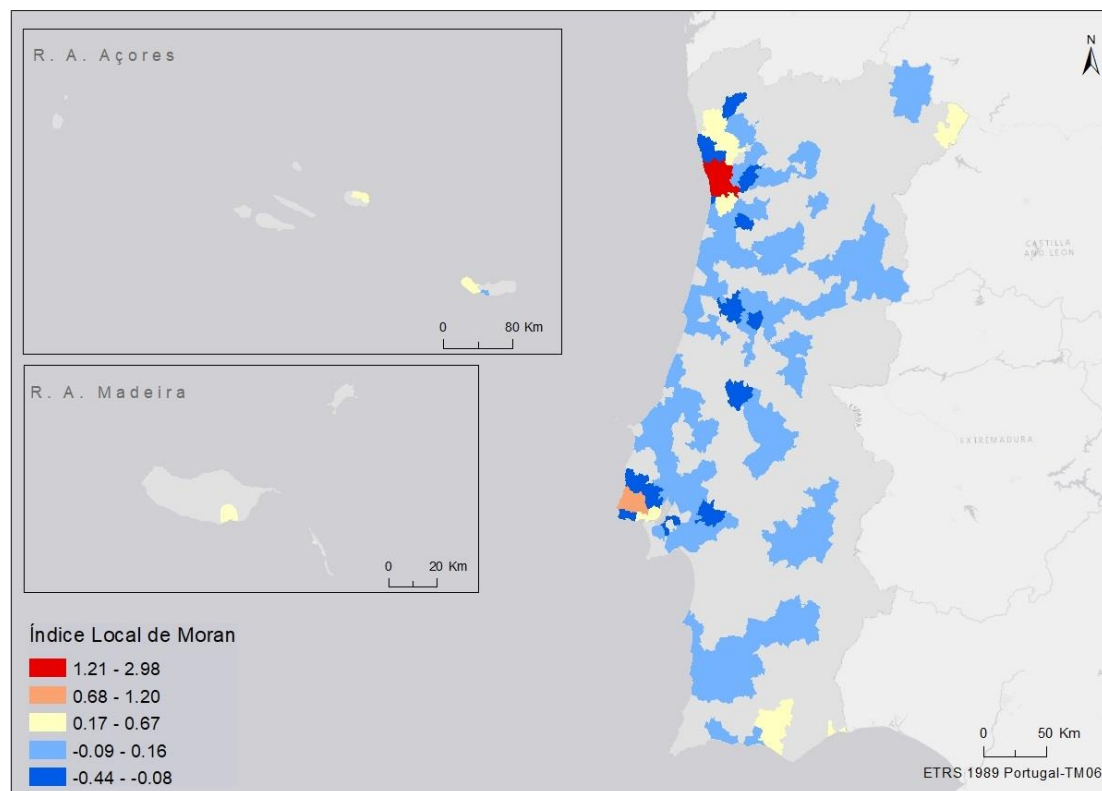


Figura 5.1. *Clusters e outliers*²³ espaciais dos dadores efetivos em Portugal.

Os *clusters* com valores de índice positivo, estão localizados em municípios onde a concentração de dadores é mais intensa (figura 5.6), assim como, onde a população dos 45 aos 65 anos de idade é mais elevada (idade média dos dadores), como é possível constatar através da figura 5.10.

5.1.2. Regressão Linear Múltipla

De acordo com a informação disponibilizada pelo Instituto Nacional de Estatística, foram selecionadas 15 variáveis (Quadro 5.1), de modo a serem utilizadas como fator explicativo da aglomeração geográfica de dadores (variável dependente).

²³ Um *outlier* espacial refere-se a uma localidade cujo valor é elevado e encontra-se rodeado por valores baixos ou vice-versa (Anselin *et al.*, 2010; Longley *et al.*, 2005; Waller e Gotway, 2004)

Quadro 5.1. Variáveis explicativas pré-selecionadas

Código	Variáveis independentes
V1	Óbitos por Acidente de Viação
V2	Óbitos por Doença do Aparelho Circulatório
V4	População Residente entre os 45 aos 69 anos de idade
V5	Número de Indivíduos Residentes
V6	Número de Urgências Hospitalares por Município
V7	Óbitos por Quedas Acidentais
V8	Óbitos por Suicídio
V9	Óbitos por Homicídios e Lesões
V10	Ocorrências Hospitalares por Acidente de Viação
V11	Ocorrências Hospitalares por Agressão
V12	Ocorrências Hospitalares por Hemorragia
V13	Ocorrência Hospitalar por Obstrução da Via Aérea
V14	Ocorrência Hospitalar por Paragem Cardiorrespiratória
V15	Ocorrência Hospitalar por Trauma

Escolhidas as variáveis, procedeu-se à normalização das mesmas, de modo a que as diferenças nas unidades não influenciem os resultados. Através das possíveis dependências, dá-se a presença da multicolinearidade, onde por meio da técnica VIF se identificaram e descartaram as variáveis explicativas redundantes. No Quadro 5.2 é possível verificar a contabilização dos valores de VIF para cada uma das variáveis pré-selecionadas. Sendo assim, as variáveis relevantes a considerar são o número de: óbitos por acidente de viação; óbitos por doença do aparelho circulatório; óbitos por quedas acidentais; óbitos por suicídio e óbitos por homicídio e lesões.

De entre todas as variáveis independentes em estudo, a variável óbitos por doença do aparelho circulatório é a que apresenta um melhor valor de VIF, o que vem comprovar a sua fiabilidade, dado que esta é uma das principais causas de morte em Portugal, assim como a que mais contribui para a doação de órgãos após a morte. Quanto às restantes variáveis selecionadas, estas encaixam-se muito possivelmente nos óbitos por “TCE-

Outro”²⁴, sendo esta a segunda principal fonte de dados, como é possível constar através da figura 4.14.

Quadro 5.2. Valores de VIF para as variáveis independentes.

Código	Variável Independente (Fator)	VIF
V1	Óbitos por Acidente de Viação	1,139
V2	Óbitos por Acidente Vascular Cerebral	57,010
V3	Óbitos por Doença do Aparelho Circulatório	1.013
V4	População Residente entre os 45 aos 69 anos de idade	17,127
V7	Óbitos por Quedas Acidentais	3,368
V8	Óbitos por Suicídio	5,952
V9	Óbitos por Homicídios e Lesões	2,668
V10	Ocorrências Hospitalares por Acidente de Viação	71,509
V11	Ocorrências Hospitalares por Agressão	26,779
V12	Ocorrências Hospitalares por Hemorragia	92,340
V13	Ocorrência Hospitalar por Obstrução da Via Aérea	28,585
V14	Ocorrência Hospitalar por Paragem Cardiorrespiratória	75,460

Posto isto, foi realizada uma análise exploratória em que foram corridos modelos. Foram testadas 270 combinações das quais apresentam um R^2 ajustado da regressão de 0,75, o que significa que as variáveis independentes escolhidas explicam 75% da variável dependente. Este coeficiente de determinação (R^2), avalia a proporção da variância da variável dependente, que por sua vez é explicada pelas variáveis independentes (Saldiva e Alvarenga, 2003). Se o valor de R^2 for igual a 1 ($R=1$), indica a existência de uma relação linear exata entre os *outputs* e os alvos, porém se o valor de R^2 for próximo de zero expressa a ausência de relação linear entre os dados (Demuth *et al.*, 2015).

Na prática (Figura 5.2) face às condições observadas, a informação acerca dos resíduos indica a existência de municípios onde o número de dados é superior ao expectável (resíduos positivos), em oposição, a existência de municípios onde o número de dados

²⁴ Entenda-se como outras causas de óbito por Traumatismo Crânio Encefálico, e abrange os óbitos por quedas acidentais, suicídios e por homicídios e lesões.

se encontra abaixo do seu potencial (resíduos negativos). Dos municípios com resíduos positivos elevados ($> 2,5$) destacam-se por ordem decrescente o município de Sintra, Gondomar, Valongo, Maia, Matosinhos, Santo Tirso, Amadora e Porto. Face a resíduos abaixo do expectável ($< -2,5$) os municípios são: Lisboa, Almada, Viana do Castelo, Torres Vedras, Castelo Branco e Leiria.

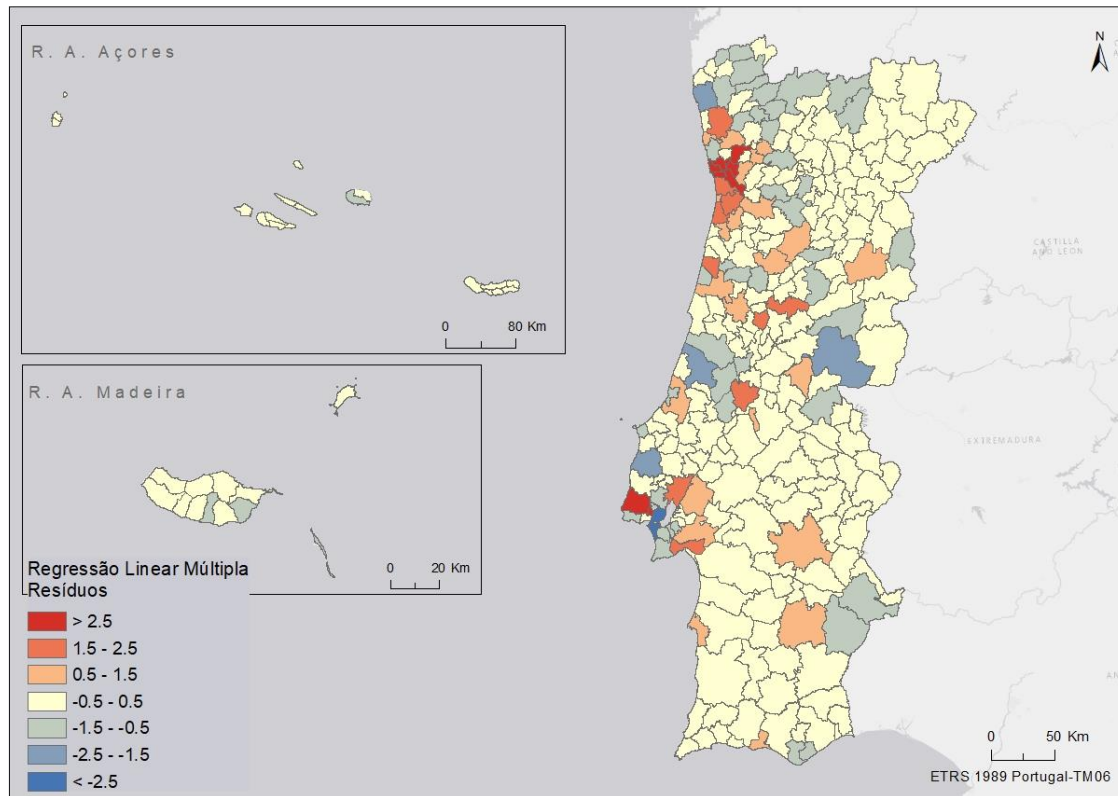


Figura 5.2. Resíduos da regressão linear múltipla

Municípios abaixo do expectável são áreas que necessitam de especial atenção dado o seu potencial de doação. É de relevância levar a cabo ações de sensibilização nesses hospitais, fazendo chegar às pessoas mais informação relativamente ao tema, assim como, exercer uma redobrada atenção por parte dos coordenadores hospitalares de doação, na deteção de possíveis dadores.

5.1.3. Árvores de classificação e regressão

A abordagem geral para derivar predições a partir de algumas condições simples do tipo se-então (*if-then*), também pode ser aplicada a problemas de regressão. A questão de base foi determinar os correspondentes dos dadores, *i.e.*, as variáveis que melhor predizem a o número de dadores de órgãos num determinado concelho. Uma análise dos dados, recorrendo a uma árvore de regressão permitiu determinar a melhor árvore (Figura 5.3).

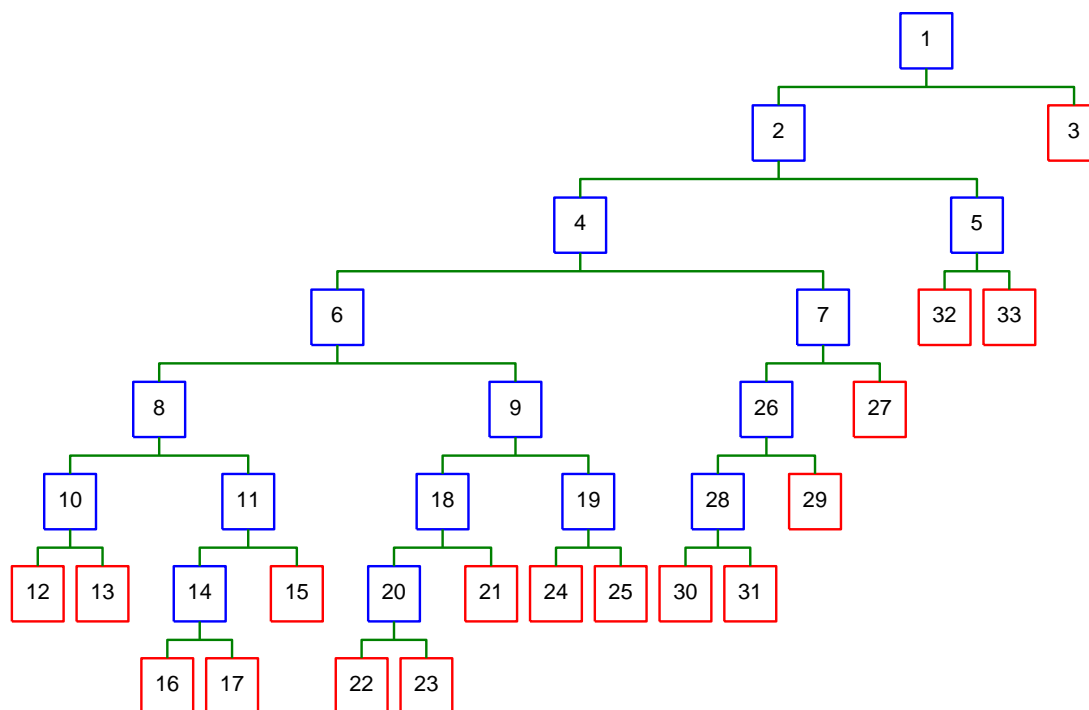


Figura 5.3. Árvore de regressão para os fatores explicativos dos dadores efetivos

A interpretação destes resultados (Figura 5.3) é bastante direta (Quadro 5.3): os municípios em que o valor normalizado de Ocorrências Hospitalares por Trauma é inferior 1,718692 possuem menos dadores que aqueles cujo valor é superior. Estes resultados são simples, fáceis de apresentar e intuitivamente claros.

Quadro 5.3. Parâmetros (estrutura simplificada) de divisão da árvore de regressão dos dados efetivos

Nó	Divisão esquerda	Divisão direita	Dimensão	Fator divisão	Valor divisão
1	2	3	305	Ocorrências Hospitalares por Trauma	1.718692
2	4	5	294	Óbitos por Doença do Aparelho Circulatório	0.173898
4	6	7	244	Óbitos por Doença do Aparelho Circulatório	-0.249207
6	8	9	162	Óbitos por Doença do Aparelho Circulatório	-0.399902
8	10	11	89	Óbitos por Doença do Aparelho Circulatório	-0.475250
18	20	21	42	Óbitos por Doença do Aparelho Circulatório	-0.281085
20	22	23	37	Óbitos por Doença do Aparelho Circulatório	-0.365126
19	24	25	31	Óbitos por Doença do Aparelho Circulatório	-0.312963
7	26	27	82	Óbitos por Doença do Aparelho Circulatório	-0.017369
5	32	33	50	Óbitos por Doença do Aparelho Circulatório	0.683942

Selecionado o modelo de árvore são identificadas as variáveis preditoras mais importantes (Quadro 5.4). Neste caso, optou-se ou isolar as quatro principais como as que mais contribuem para a caracterização do perfil dos municípios quanto ao número de dados.

Quadro 5.4. Variáveis preditoras (independentes) a utilizar no modelo e respetiva importância.

	Variável	Importância (peso)
Óbitos por Doença do Aparelho Circulatório	100	1,000
Óbitos por Acidente Vascular Cerebral	98	0,978
Ocorrências Hospitalares por Trauma	96	0,964
Ocorrências Hospitalares por paragem Cardiorrespiratória	96	0,959
Ocorrências Hospitalares por Hemorragia	96	0,955
Ocorrências Hospitalares por Acidente de Viação	92	0,917
Ocorrências Hospitalares por Agressão	92	0,916
Ocorrências Hospitalares por Obstrução das Vias Aéreas	91	0,907
Óbitos por Suicídio	80	0,797
Óbitos por Quedas Acidentais	80	0,797
Óbitos por Homicídios e Lesões	64	0,643
Óbitos por Acidente de Viação	19	0,187
Ocorrências Hospitalares por Doença do Aparelho Circulatório	2	0,021

5.1.4. Redes Neurais Artificiais

Atualmente, são inúmeras as investigações que têm demonstrado a importância do papel das RNA nos estudos aplicados à saúde, já que se trata de um modelo computacional extremamente útil na resolução de problemas clínicos complexos, assim como no apoio à decisão (Batista, *et al.*, 2014).

O *software Statistica* dispõe de vários tipos de RNA a partir da qual é possível modelar fenómenos. No presente caso de estudo, dentro da categoria das RNA recorreu-se ao modelo de rede Perceptron Multicamada (do inglês *multilayer perceptron* - MLP) e rede Linear (do inglês *linear network* – LN).

Inicialmente, a RNA admitiu um conjunto de *inputs* para cada neurónio (14 variáveis independentes). Estas são variáveis, de um modo geral, referentes às causas de morte dos doadores de órgãos falecidos. Mais especificamente, abrange variáveis de cariz populacional (idade média dos doadores), passando pelas ocorrências hospitalares, e mais fulcral, os óbitos. Cada *input* foi conduzido através de uma conexão à qual foi gerado um peso, que por sua vez corresponde à eficácia sináptica do neurónio.

Para obtenção de resultados, recorreu-se ao treino da rede. O programa forneceu os resultados das cinco melhores redes, baseado nos erros de validação e erro do teste. A melhor rede é aquela que apresenta um menor erro de treino. Tais resultados incluem o perfil da rede (números de entrada, camadas escondidas e saída), níveis de acerto e apresenta o erro nas fases de treino, validação e teste. Nos campos “performance de treino”, “performance de validade” e “performance de teste” indicam a *performance* da rede face à seleção, treino e respetivos testes. O campo “perfil da rede” é fulcral uma vez que contém informação associada, nomeadamente, o tipo de rede, o número de neurónios em cada camada, e ainda o número de variáveis *input* e *output*. O campo “treino” refere-se ao algoritmo utilizado para treinar a rede.

Na prática e segundo o Quadro 5.5, o perfil da rede utilizado no estudo é MLP sendo o resultado obtido designado por “BP100, CG20”, i.e., expressa que foram utilizadas 100

épocas de treino através do algoritmo *Back Propagation* seguido de 20 épocas com o algoritmo *conjugate gradient descent*, onde *a posteriori* o treino foi parado devido à sob aprendizagem, regressando assim ao início (retro propagação do erro). Uma outra rede utilizada no estudo diz respeito à rede Linear, tendo esta por algoritmo de treino a designação “PI” (do inglês *Pseudo Inverse*). Do conjunto de 5 melhores resultados destacam-se as três redes Lineares, com valores de erro de teste que vão dos 0,30 aos 0,31. Porém, face ao erro de treino a rede que apresentou um melhor resultado é a MLP (5ª) dado que obteve um erro de treino de “0,011091”, revelando-se este o menor resultado e, portanto, a rede com menor erro.

Quadro 5.5. Análise da *performance* das cinco melhores Redes Neurais Artificiais para os diferentes conjuntos de variáveis.

Relatório de Resumo do Modelo (Variáveis Dadores de Órgãos)											
Nº	Perfil da Rede	Performance Treino	Performance Validade	Performance Teste	Erro de Treino	Erro de Seleção	Erro de Teste	Treino	Input	Neurónios Escondidos (1)	Neurónios Escondidos (2)
1	Linear	0.091497	0.201420	0.106511	0.013375	0.020972	0.031478	PI	13	0	0
2	Linear	0.090225	0.199430	0.106258	0.013189	0.020784	0.031326	PI	14	0	0
3	Linear	0.092783	0.195075	0.104035	0.013563	0.020275	0.030801	PI	12	0	0
4	MLP	0.084611	0.153415	0.255710	0.012369	0.015870	0.074135	BP100, CG20, C G61b	8	10	4
5	MLP	0.075852	0.150344	0.311246	0.011091	0.015580	0.090334	BP100, CG20, C G47b	5	2	0

Com a finalidade de avaliar a capacidade preditiva e a validade do modelo, foi realizada uma análise de sensibilidade. Esta fornece uma visão da utilidade e importância das variáveis independentes na explicação da variável dependente (dadores de órgãos falecidos), onde por sua vez o modelo acaba por produzir um *ranking*.

Deste modo e face aos resultados da análise sensitiva (Quadro 5.6), a variável “óbitos por doença do aparelho circulatório” encontra-se como uma das mais relevantes, na justificação da localização dos dadores efetivos falecidos. Esta ocupa o primeiro lugar no *ranking* das 13 variáveis independentes, seguindo-se de um segundo lugar as “ocorrências hospitalares por hemorragia”. Denote-se que as doenças do aparelho circulatório são a principal causa de morte em Portugal também a principal causa dos óbitos associados aos dadores falecidos efetivos. Estas afirmações validam os resultados da análise de sensibilidade das árvores de decisão já que nessa análise o primeiro lugar do *ranking* também é ocupado pelas doenças do aparelho circulatório.

Quadro 5.6. Análise sensitiva das variáveis.

Variáveis	Relação	Ranking
Óbitos por Doença do Aparelho Circulatório	5,340	1
Ocorrências Hospitalares por Hemorragia	2,983	2
Ocorrências Hospitalares por Trauma	1,664	3
Óbitos por Acidente de Viação	1,006	4
Óbitos por Suicídio	1,003	5
Ocorrências Hospitalares por Doença do Aparelho Circulatório	0,994	6
Ocorrências Hospitalares por Acidente de Viação	0,982	7
Ocorrências Hospitalares por paragem Cardiorrespiratória	0,953	8
Óbitos por Quedas Acidentais	0,922	9
Ocorrências Hospitalares por Obstrução das Vias Aéreas	0,918	10
Óbitos por Acidente Vascular Cerebral	0,913	11
Ocorrências Hospitalares por Agressão	0,870	12

Num terceiro, quarto e quinto lugar a análise sensitiva destaca as ocorrências hospitalares por trauma, os óbitos por acidente de viação e os óbitos por suicídio, respetivamente. Estes estão associados aos traumatismos crânio encefálicos, que são a segunda principal causa dos óbitos dos dadores efetivos falecidos em Portugal.

A utilização das RNA como método mais recente e atualmente em voga, quando comparado ao método Regressão Linear Múltipla, é um método mais fechado dado que os pesos são atribuídos automaticamente através dos neurónios. Porém, a RLM embora

seja um método mais tradicional, possui uma maior abertura, uma vez que proporciona ao utilizador a possibilidade de seleção das variáveis tendo em conta o erro produzido, perante o assunto estudado. Uma das desvantagens das redes neuronais, passa por serem sistemas *black box* – caixa preta, i.e., a representação interna dos dados no processador (a rede) não é revelada (Buch, 2000).

Confrontando os resultados dois métodos (RLM e RNA) a variável “óbitos por doença do aparelho circulatório” encontra-se em ambos os casos como a mais preponderante na explicação da localização dos dados, seguindo-se os óbitos por suicídio e os óbitos por acidente de viação.

5.2. Doentes Transplantados

5.2.1. Índice Global de Moran

Através da distância euclidiana e do método *Inverse Distance*, o índice global de Moran foi aplicado às várias tipologias de doentes transplantados. Este índice foi calculado para os municípios, utilizando o número de doentes transplantados. Analisando os valores (Quadro 5.7) é possível constatar que em todas as análises, o índice global de Moran's foi ligeiramente superior a 0, ou seja, ficou acima do índice esperado indicando uma autocorrelação positiva. Quanto ao valor de *z-score*, este é superior que 2,58 em todas as variáveis, o que também reforça a autocorrelação positiva. Neste destacam-se os doentes transplantados de rim, seguindo-se dos doentes transplantados de fígado. Relativamente ao valor de *p-value* este é bastante diminuto, o que demonstra a impossibilidade de o padrão espacial observado tenha sido gerado por um padrão aleatório, sendo então possível rejeitar a hipótese nula (i.e., independência espacial), dado a existência de significância estatística nas observações.

Quadro 5.7. Índice Global de Moran referente às várias tipologias de doentes transplantados em Portugal.

Atributo Analisado	Índice de Moran	Índice esperado	z-score	p-value
Transplantados Órgão: Fígado	0,078	-0,003	14,355	0,000000
Transplantados Órgão: Pulmão	0,034	-0,003	7,238	0,000000
Transplantados Órgão: Pâncreas	0,029	-0,003	5,920	0,000000
Transplantados Órgão: Rim	0,106	-0,003	19,777	0,000000
Transplantados Órgão: Coração	0,023	-0,003	4,688	0,000003

5.2.2. Análise de Hot Spot (Getis-Ord G_i^*)

Através da análise *hot spot* analisamos a ocorrência de aglomerados espaciais de doentes transplantados nas suas várias tipologias de órgão. Realizada a análise ao município, subsistem apenas alguns valores de *hot spot* (pontos quentes), dado que na sua envolvente não existem casos suficientemente “robustos” de pessoas transplantadas, o que leva a que na maioria dos casos (municípios) sejam considerados insignificantes.

Face a resultados referente a doentes transplantados com rim (Figura 5.4), o primeiro aspeto a salientar é a disposição de vários aglomerados de *hot spot* concentrados em 11 municípios estatisticamente significativos, todos eles localizados no litoral do país.

Os *clusters* espaciais com uma maior significância (99%), localizam-se nos seguintes municípios: Coimbra, Cascais e Lisboa, seguindo-se com 95% de significância Guimarães, Montijo, Setúbal, Oeiras, Amadora, e com 90% o Funchal, Porto e Matosinhos. Dado que o rim é o órgão mais transplantado em Portugal, e estes são municípios com elevado cariz populacional, tal facto pode explicar a presença de pontos quentes nestas áreas geográficas.

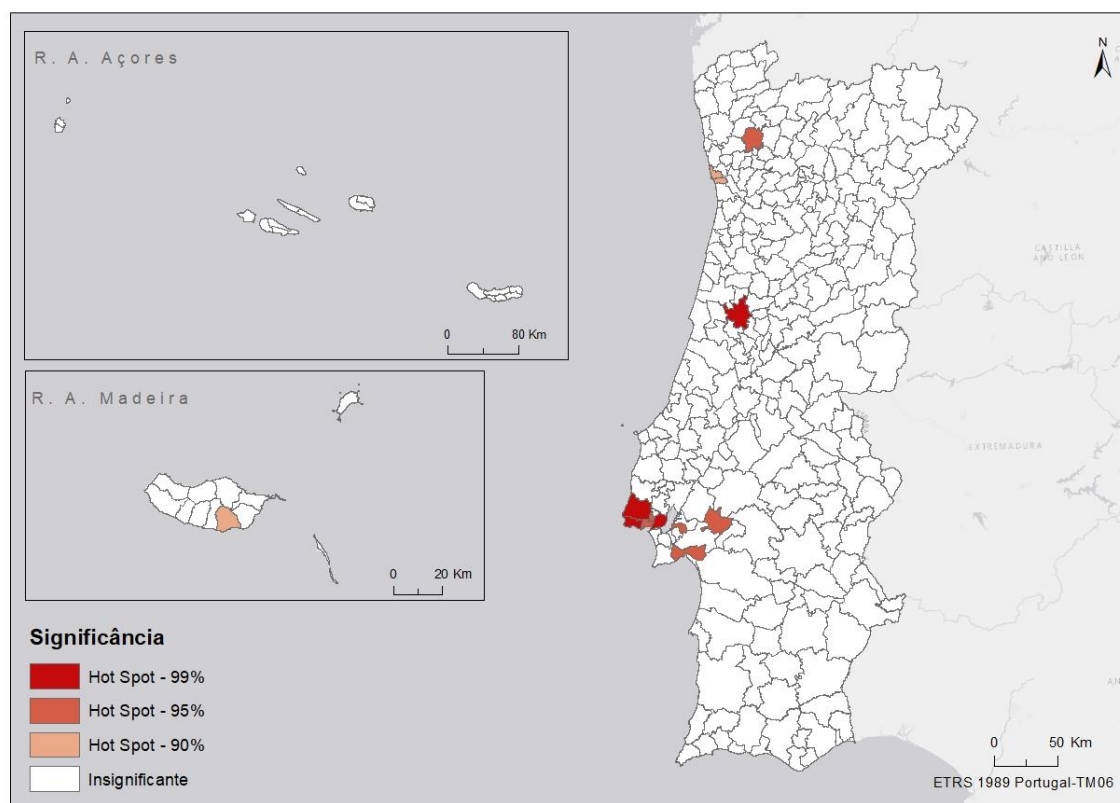


Figura 5.4. *Clusters* espaciais de doentes transplantados de rim nos municípios de Portugal

Relativamente ao resultado da análise *hot spot* de doentes transplantados de fígado (Figura 5.5), o segundo órgão mais transplantado em Portugal, também a grande maioria dos municípios estão localizados no litoral do país. São 16 os municípios estatisticamente significativos, com uma significância que subsiste entre os 90 e os 99%. Assim, os *clusters* espaciais com uma maior significância (99%) localizam-se nos seguintes municípios: Porto, Figueira da Foz, Coimbra, Cascais, Sintra, Amadora e Lisboa, com significância de 95% o município de Ílhavo, e com 90% destaque para Braga, Valongo, Matosinhos, Vila Nova de Gaia, Viseu, Carregal do Sal, Loures e Seixal. Nas ilhas constata-se uma ausência e insignificância de *clusters* alusivos à transplantação hepática.

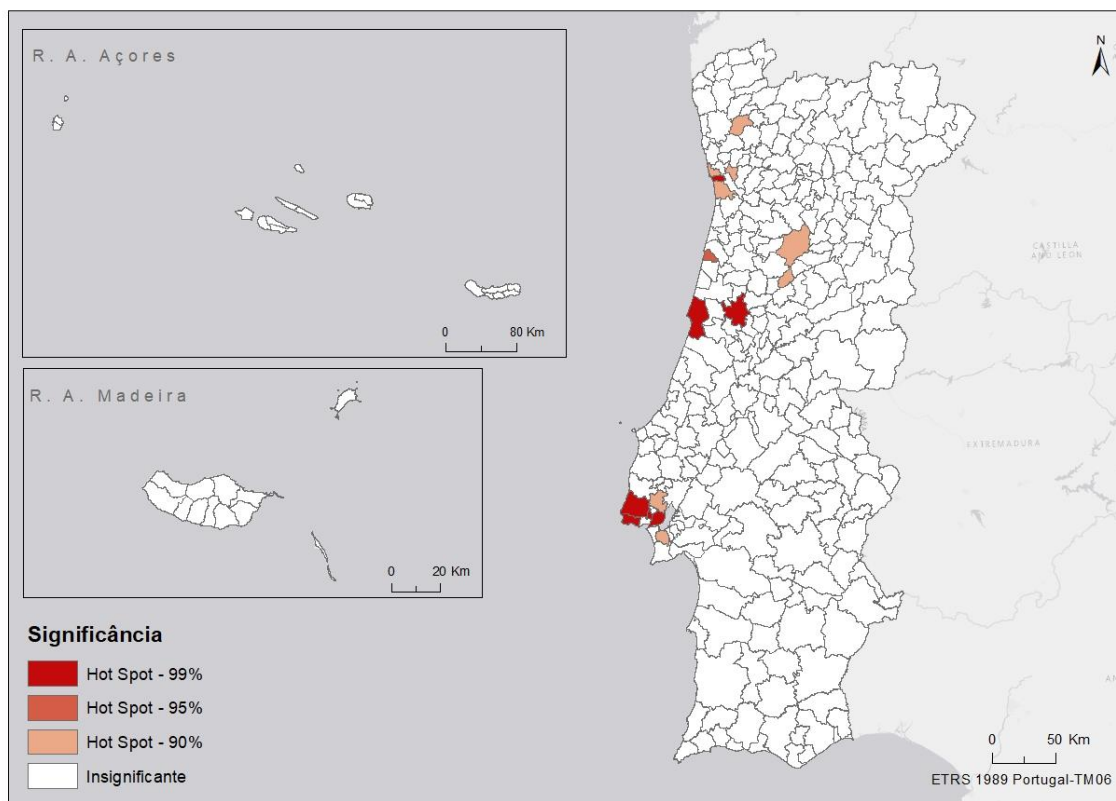


Figura 5.5. *Clusters* espaciais de doentes transplantados de fígado nos municípios de Portugal

No que respeita ao resultado da análise *hot spot* de doentes transplantados de pulmão (Figura 5.6), os pontos quentes espaciais de maior magnitude têm por foco municípios pertencentes às duas grandes metrópoles, ou seja, onde se concentra o maior conglomerado populacional.

São 12 os municípios estatisticamente significativos, com um intervalo que varia entre os 95% e os 99%. Com uma significância de 99 % estão os municípios de Sintra, Cascais, Vila Nova de Gaia e Valongo, e com uma significância de 95% Bragança, Almeida, Santo Tirso, Porto, Ílhavo, Lisboa, Seixal e Lagos.

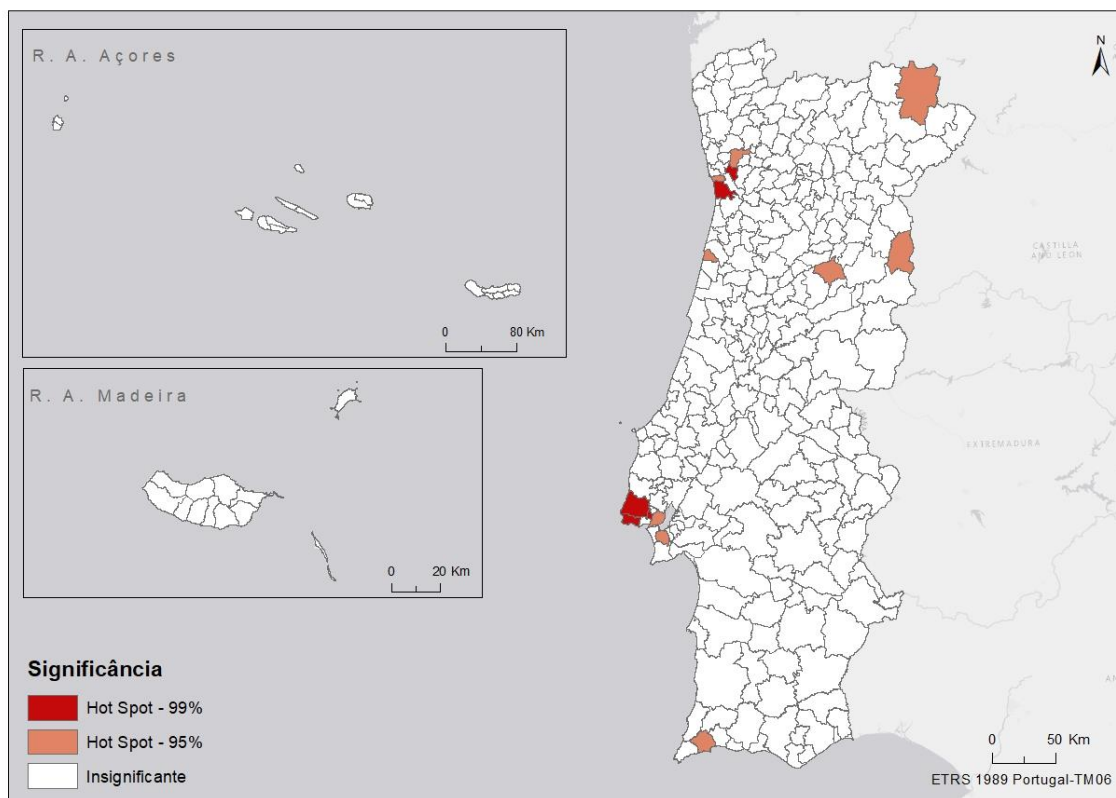


Figura 5.6. *Clusters* espaciais de doentes transplantados de pulmão nos municípios de Portugal.

Observando o resultado da análise *hot spot* referente a doentes cardíacos transplantados (Figura 5.7), um dos primeiros aspetos a salientar é a ocorrência de pontos quentes em municípios pertencentes ao interior Norte e Centro do continente. Como analisado anteriormente (Figura 4.20), existe uma resposta positiva do interior face à transplantação, porém a presente significância vem demonstrar que grande parte desses doentes são os transplantados de coração.

Um segundo aspeto está relacionado com a existência de valores quentes que vão dos 95% aos 99%. Assim, os *clusters* espaciais com uma maior significância (99%) localizam-se nos seguintes municípios: Póvoa de Varzim, Matosinhos, Vila Nova de Gaia, Oliveira do Bairro, Coimbra, Lousã, Pombal, Cascais, Amadora, Lisboa e com uma significância de 95% encontram-se municípios pertencentes aos arquipélagos, como é o caso de São Roque do Pico e Funchal.

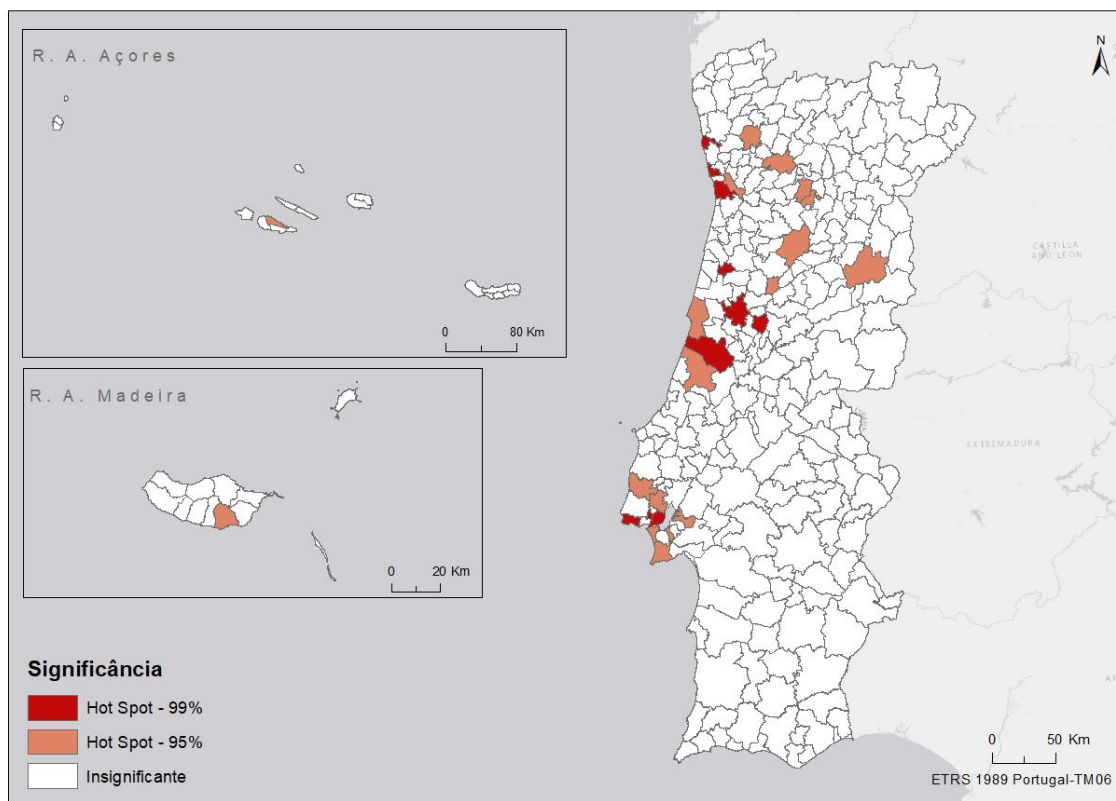


Figura 5.7. *Clusters* espaciais de doentes transplantados de coração nos municípios de Portugal.

O número de doentes por transplante pancreático (Figura 5.8) representa somente *clusters* com significância de 99%. Tal facto deve-se ao diminuto número de doentes transplantados no global e, conseqüentemente, por município (varia entre 1 e 3 doentes transplantados), logo a presença de padrões será muito homogénea.

Sendo assim, os municípios estatisticamente significativos são os municípios de: Vila Pouca de Aguiar, Vila Nova de Famalicão; Matosinhos, Porto, Vila Nova de Gaia, Oliveira de Azeméis, Soure, Cartaxo, Sintra, Cascais, Lisboa, Almada, Seixal e Montijo.

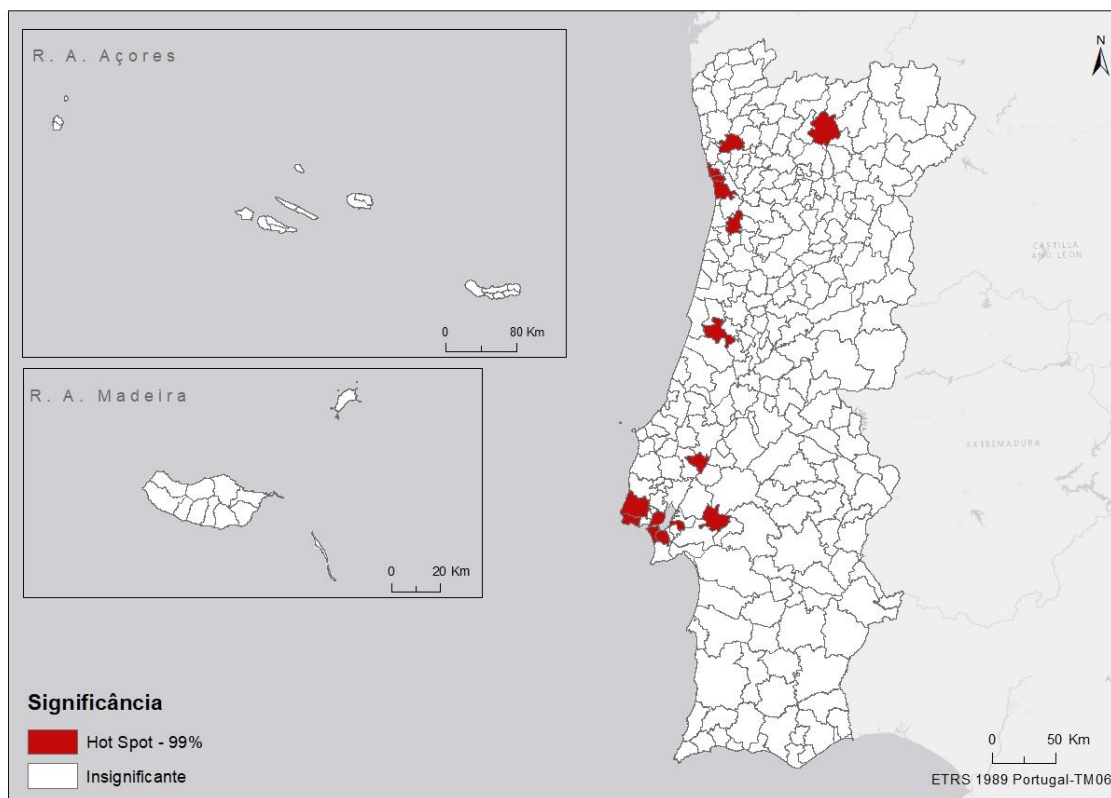


Figura 5.8. Clusters espaciais de doentes transplantados de pâncreas nos municípios de Portugal.

5.2.3. Regressão Linear Múltipla

De acordo com a informação disponibilizada pelo INE, foram selecionadas 11 variáveis (Quadro 5.8), que representam patologias que poderão levar à falência de órgãos e à necessidade de transplante e, como tal poderão ter influência na explicação do número de doentes transplantados (variável dependente).

Escolhidas as variáveis procedeu-se à normalização das mesmas, de modo a que as diferenças nas unidades não influenciem os resultados. Face à presença de multicolinearidade, através da técnica *VIF* descartaram-se as variáveis explicativas redundantes. No Quadro 5.9 é possível verificar a contabilização dos valores de *VIF* para cada uma das variáveis pré-selecionadas. Sendo assim, as variáveis relevantes a considerar são: os óbitos por doença isquémica do coração; óbitos por outras doenças cardíacas; óbitos por doença crónica do fígado; óbitos por malformações congénitas e

anomalias; óbitos por malformações congênitas do aparelho circulatório e a população residente que usufrui do rendimento social de inserção. Esta última variável foi considerada uma vez que rendimentos mais baixos estão normalmente associados a piores cuidados de saúde. “*Várias investigações revelam que, independentemente do nível de análise considerado, do local ao nacional, maiores desigualdades nos rendimentos associam-se a piores resultados na saúde*” (Nogueira, 2008).

Quadro 5.8. Variáveis explicativas pré-selecionadas.

Código	Variável independente
V1	Óbitos por Diabetes <i>Mellitus</i>
V2	Óbitos por Doença do Aparelho Circulatório
V3	Óbitos por Doença do Aparelho Respiratório
V4	Óbitos por Doença Isquémica do Coração
V5	Óbitos por outras Doenças Cardíacas
V6	Óbitos por Doença Crónica do Fígado
V7	Óbitos por Doença do Rim e Ureter
V8	Óbitos por Malformações congênitas e anomalias
V9	Óbitos por Malformações congênitas do Aparelho Circulatório
V10	População que usufrui do Rendimento Social de Inserção
V11	População Residente dos 65 aos 85 anos de idade

Quadro 5.9. Valores de VIF das variáveis independentes.

Código	Variável Independente (Fator)	VIF
V1	Óbitos por Diabetes <i>Mellitus</i>	15,948
V3	Óbitos por Doença do Aparelho Respiratório	25,493
V4	Óbitos por Doença Isquémica do Coração	9,972
V5	Óbitos por outras Doenças Cardíacas	9,288
V7	Óbitos por Doença Crónica do Fígado	5,750
V8	Óbitos por Doença do Rim e Ureter	11,218
V9	Óbitos por Malformações congênitas e anomalias	9,333
V10	Óbitos por Malformações congênitas do Aparelho Circulatório	5,122
V11	População que usufrui do Rendimento Social de Inserção	4,598
V12	População Residente dos 65 aos 85 anos de idade	43,882

Posto isto, foi realizada uma análise exploratória em que foram corridos modelos. Foram testadas 306 combinações das quais apresentam um R^2 ajustado da regressão de 0,89, o que significa que as variáveis dependentes escolhidas explicam 89% da variável dependente.

Na prática, analisando a Figura 5.9 face às condições observadas, a informação acerca dos resíduos indica a existência de municípios onde o número de doentes transplantados é superior ao expectável (resíduos positivos), e vice-versa (resíduos negativos). Dos municípios com resíduos positivos elevados ($> 2,5$) destaca-se por ordem decrescente o município de Sintra, Coimbra, Valongo, Cascais e Montijo. Municípios com resíduos negativos elevados representam os locais onde face às patologias seleccionadas no estudo, haverá a possibilidade de o número de doentes transplantados aumentar. Por ordem decrescente ($< 2,5$) estes municípios são: Lisboa, Santa Maria da Feira, Funchal e Almada.

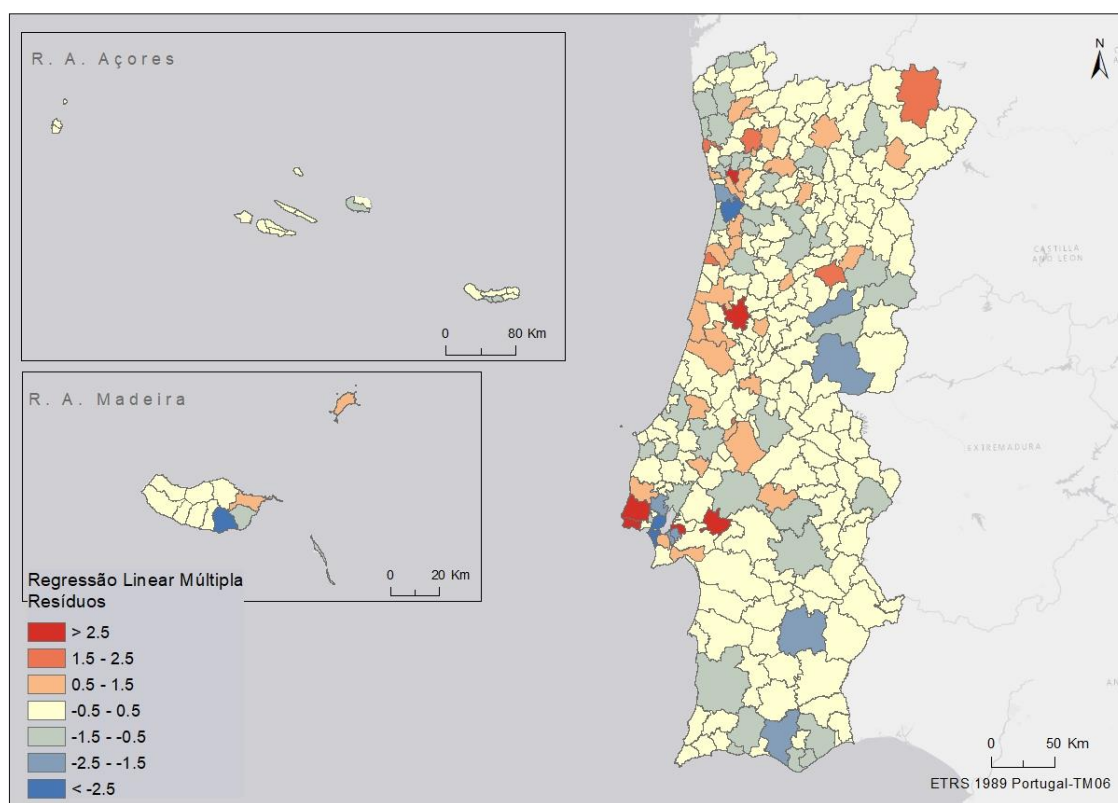


Figura 5.9. Resíduos da regressão linear múltipla.

Os municípios onde os números de doentes transplantados se encontrem abaixo do que seria conjecturável, são áreas onde se deve intervir, de modo a prevenir e compreender o porquê do agravamento de certas patologias em determinados territórios.

5.2.4. Árvores de classificação e regressão

Também nesta fase as Árvores de Classificação e Regressão (CRT) foram usadas para desenvolver modelos que podem classificar os municípios em várias categorias de risco (=necessidade de um transplante). Cada variável independente é examinada e é criada uma divisão para maximizar a sensibilidade da classificação, resultando numa árvore de decisão (Figura 5.10). O objetivo da é desenvolver uma árvore com o tamanho ideal e a menor taxa de erro. O método CRT é capaz de determinar interações complexas entre variáveis na árvore final, em contraste com a identificação e definição das interações num modelo de regressão logística multivariada.

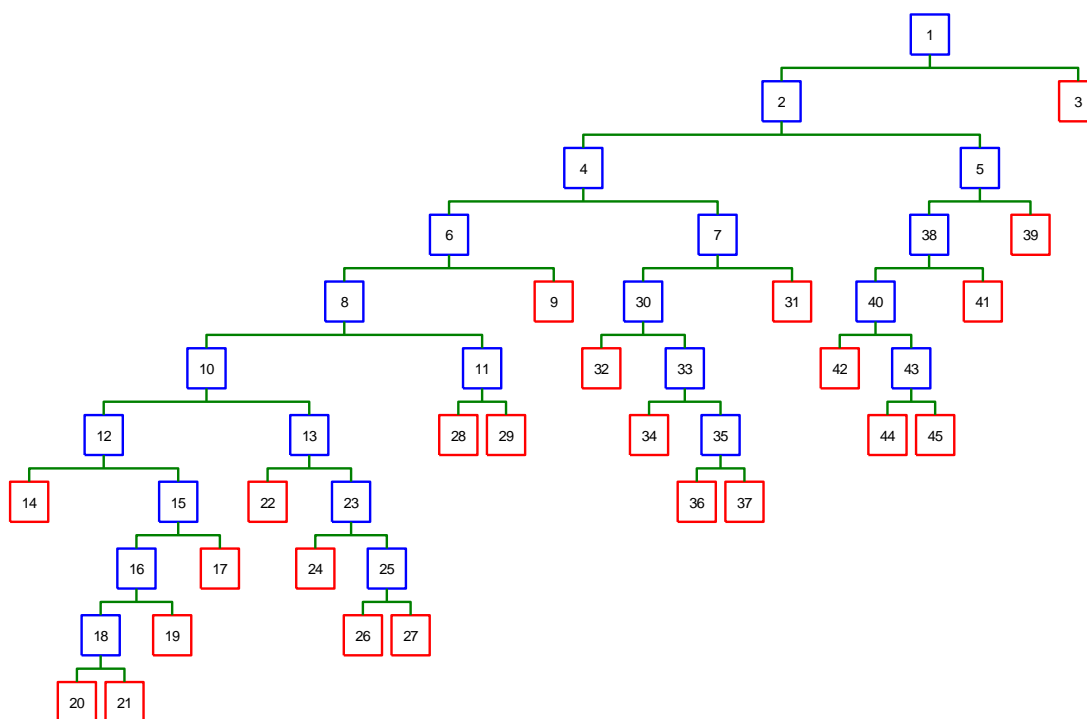


Figura 5.10. Árvore de regressão para os fatores explicativos dos doentes transplantados

A interpretação da árvore segue o mesmo esquema definido para o Quadro 5.3. Neste caso (Quadro 5.10), os Óbitos por Acidente Vascular Cerebral acima, ou abaixo, do valor normalizado de 1,998889 (= 169 óbitos) explicam a primeira divisão na caracterização dos municípios face ao número de doentes transplantados.

Quadro 5.10. Parâmetros (estrutura simplificada) de divisão da árvore de regressão dos doentes transplantados

Nó	Divisão esquerda	Divisão direita	Dimensão	Fator divisão	Valor divisão
1	2	3	306	Óbitos por Acidente Vascular Cerebral	1,998889
6	8	9	181	Óbitos por Diabetes <i>Mellitus</i>	0,038276
12	14	15	97	Óbitos por Doença do Aparelho Respiratório	-0,583996
16	18	19	53	Óbitos por Acidente Vascular Cerebral	-0,197948
18	20	21	51	Óbitos por Doença do Aparelho Respiratório	-0,510075
13	22	23	49	Óbitos por Doença do Aparelho Respiratório	-0,396816
11	28	29	32	Óbitos por Doença do Aparelho Respiratório	-0,361967
7	30	31	64	Óbitos por Acidente Vascular Cerebral	-0,023716
33	34	35	34	Óbitos por Doença Isquémica do Coração	-0,338544
35	36	37	33	Óbitos por Doença Isquémica do Coração	-0,095729
38	40	41	46	Óbitos por Acidente Vascular Cerebral	1,067127
40	42	43	37	Óbitos por Doença do Rim e Ureter	-0,413458
43	44	45	36	Óbitos por Diabetes <i>Mellitus</i>	0,453286

Após a análise da CRT, são identificadas e ordenadas as variáveis preditoras do modelo (Quadro 5.11). Mais uma vez se optou por destacar as quatro que mais contribuem para a diferenciação.

Quadro 5.11. Variáveis preditoras (independentes) a utilizar no modelo e respetiva importância.

	Variável	Importância (peso)
Óbitos por Acidente Vascular Cerebral	100	1,000
Óbitos por Doença do Aparelho Circulatório	96	0,964
Óbitos por Diabetes Mellitus	90	0,900
Óbitos por Doença do Rim e Ureter	88	0,882
Óbitos por Doença do Aparelho Respiratório	87	0,874
População que usufrui do Rendimento Social de Inserção	85	0,855
Óbitos por outras Doenças Cardíacas	82	0,823
Óbitos por Doença Isquémica do Coração	81	0,814
Óbitos por Malformações Congénitas e Anomalias	71	0,709
Óbitos por Malf. Congén. do Aparelho Circulatório	69	0,686
Óbitos por Doença Crónica do Fígado	66	0,660

5.2.5. Redes Neurais Artificiais

Inicialmente, a RNA admitiu um conjunto de *inputs* para cada neurónio (12 variáveis independentes). Estas são variáveis que representam patologias que poderão levar à falência de órgãos e à necessidade de transplante. Tais variáveis estão relacionadas maioritariamente com os óbitos, porém, uma pequena porção diz respeito a variáveis de cariz populacional, também estas possivelmente relacionadas a quem necessita de um órgão. Cada *input* foi conduzido através de uma conexão à qual foi gerado um peso, que por sua vez corresponde à eficácia sináptica do neurónio.

Para obtenção de resultados, recorreu-se ao treino da rede. O programa forneceu os resultados das cinco melhores redes, baseado nos erros de validação e erro do teste. O campo “perfil da rede” é fulcral uma vez que contém informação associada, nomeadamente, o tipo de rede, o número de neurónios em cada camada, e ainda o número de variáveis *input* e *output*.

Na prática e segundo o Quadro 5.12, o perfil da rede utilizado no estudo é MLP sendo o resultado obtido designado por “BP100, CG20, CG0b”, i.e., expressa que foram utilizadas 100 épocas de treino através do algoritmo *Back Propagation* seguido de 20 épocas com o algoritmo *conjugate gradient descent*, onde *a posteriori* o treino foi parado devido à sob aprendizagem, regressando assim ao início (retro propagação do erro). Do conjunto de 5 melhores resultados destaca-se a última rede (5ª) com valores de erro de teste de 1,71. Porém, face ao erro de treino a rede que apresentou um melhor resultado é a primeira rede (MLP) dado um erro de treino de “1,377773” revelando-se este o menor resultado e, portanto, a rede com menor erro.

Quadro 5.12. Análise da *performance* das cinco melhores RNA para os diferentes conjuntos de variáveis.

Relatório de Resumo do Modelo (Variáveis Doentes Transplantados)											
Nº	Perfil da Rede	Perfor- mance Treino	Perfor- mance Validade	Perfor- mance Teste	Erro de Treino	Erro de Seleção	Erro de Teste	Treino	Input	Neurónios Escondidos (1)	Neurónios Escondidos (2)
1	MLP	0,629870	0,473684	0,473684	1,377773	2,916199	1,943122	BP100, CG20, CG0b	9	13	13
2	MLP	0,623377	0,500000	0,526316	1,442639	2,822329	1,904532	BP100, CG20, CG0b	9	13	0
3	MLP	0,577922	0,526316	0,578947	1,464083	2,705588	1,720510	BP100, CG20, CG0b	9	13	0
4	MLP	0,616883	0,500000	0,513158	1,425551	2,556614	1,785484	BP100, CG20, CG0b	9	13	0
5	MLP	0,538961	0,486842	0,513158	1,524574	2,486898	1,714403	BP100, CG20, CG1b	7	13	0

Com a finalidade de avaliar a capacidade preditiva e a validade do modelo, foi realizada uma análise de sensibilidade. Esta fornece uma visão da utilidade e importância das variáveis independentes na explicação da variável dependente (doentes transplantados), onde por sua vez o modelo acaba por produzir um *ranking*.

Deste modo e face aos resultados (Quadro 5.13), a variável “óbitos por diabetes *mellitus*” encontra-se como uma das mais relevantes na explicação da variável dependente, ocupando o primeiro lugar no *ranking* número 1, seguindo-se em primeiro lugar no ranking número 2 os “óbitos por doença crónica do fígado”. Denote-se também a existência de padrões, nomeadamente na variável “óbitos por doença isquémica do coração” uma vez que o segundo lugar se repete em 3 dos *rankings*, onde o mesmo se sucede na variável “óbitos por doença crónica do fígado” com repetição do primeiro lugar em 2 dos *rankings*. Sendo assim, a diabetes *mellitus*, doença isquémica do coração e doença crónica do fígado segundo o modelo e face às restantes variáveis, são as patologias que mais contribuem na associação a doentes transplantados.

Quadro 5.13. Análise sensitiva das variáveis.

Variáveis	Relação	Ranking
Óbitos por Diabetes Mellitus	1	2,983
Óbitos por Acidente Vascular Cerebral	2	1,006
Óbitos por Doença do Rim e Ureter	3	0,994
Óbitos por Doença do Aparelho Circulatório	4	0,953
Óbitos por Doença Isquêmica do Coração	5	0,918
Óbitos por outras Doenças Cardíacas	7	0,870
Óbitos por Doença Crônica do Fígado	8	0,800
Óbitos por Doença do Aparelho Respiratório	9	0,790

Confrontando os resultados de dois modelos (CRT e RNA), ambos coincidiram nas variáveis identificadas como sendo as quatro mais importantes na explicação da localização dos doentes transplantados, nomeadamente: 1) Óbitos por Diabetes *Mellitus*; 2) Óbitos por Acidente Vascular Cerebral; 3) Óbitos por Doença do Rim e Ureter; e 4) Óbitos por Doença do Aparelho Circulatório.

Capítulo 6. Considerações Finais

Como desfecho deste relatório de estágio, destacam-se algumas observações e conclusões do proceder dos dadores efetivos e doentes transplantados no território.

Os resultados da análise estatística demonstram que é precisamente em municípios como Gondomar, Maia, Matosinhos, Valongo, Porto e Vila Nova de Gaia que subsiste uma elevada concentração espacial de pessoas dadoras *post mortem*, ou seja, é maior em municípios pertencentes à Área Metropolitana do Porto. Tal pode ser explicado devido ao facto de *clusters* com valores de índice positivo estarem localizados em municípios onde a população dos 45 aos 65 anos de idade é elevada, sendo esta a idade média dos dadores.

Numa perspetiva territorial, dado o fluxo de acesso dos dadores ao respetivo hospital onde foi realizada a cirurgia de remoção do(s) órgão(s), constatou-se dissemelhanças. Na grande maioria dos casos, o local de residência do dador efetivo encontra-se a uma distância próxima do hospital onde se realizou a cirurgia de remoção. Porém, apenas em casos mais esporádicos (1 dador) existe um distanciamento da morada em relação ao hospital. No caso das ilhas, há dadores que têm por residência as regiões autónomas, mas acabaram por ser dadores no continente e vice-versa.

Os hospitais do SNS (rede de medicina intensiva) são *stakeholders* com um papel direto e fulcral na doação e transplantação de órgãos em Portugal. É nestes hospitais que todos os anos dão entrada potenciais dadores. No entanto, é possível verificar a existência de hospitais com especialidade de neurocirurgia que possuem uma ausência de doação anual relativamente aos dois anos em estudo.

Numa análise ao número de camas de medicina intensiva (unidade de cuidados intensivos e intermédios) a relação preconizada como ideal entre número de dadores e número de camas seria de 1 dador por cama de UCI ao ano. Não obstante, em 2017 cerca de 70% dos hospitais portugueses com a especialidade de neurocirurgia encontravam-se aquém desta relação. Entre os municípios onde esta relação é

considerada média baixa ou baixa encontram-se o Hospital de Vila Nova de Gaia, Hospital do Barlavento Algarvio, Hospital de São Pedro de Vila Real, Centro Hospitalar Cova da Beira, Hospital de Santa Cruz, Hospital Distrital do Pombal e Hospital Distrital de Santarém. Por outro lado, também é possível detetar hospitais onde face ao número de camas de UCI existentes, o número de dadores ao ano é bastante superior. É o caso do Hospital Universitário de Coimbra, assim como o Hospital de Braga e o Hospital de São Teotónio.

A mesma realidade se sucede nos hospitais sem neurocirurgia. Com uma relação considerada muito baixa, destaca-se o Hospital Distrital de Santarém, o Hospital Divino Espírito Santo, o Hospital Distrital de Pombal, Hospital Nossa Sra. do Rosário e o Hospital de Santo André. Com uma relação considerada muito alta, surgem o Hospital de São Bernardo e o Hospital de Abrantes. Apenas dois hospitais se encontram próximos da relação preconizada como o ideal, sendo eles o Hospital Fernando da Fonseca e a Unidade Local de Saúde do Nordeste. Tais resultados dão-nos a entender que, embora em 2017 Portugal tenha ocupado um lugar de destaque na doação a nível europeu e mundial, ainda existe margem e potencial para melhoria.

Através do resultado da análise de potencial, municípios como Sintra, Gondomar, Valongo, Maia, Matosinhos, Santo Tirso, Amadora e Porto estão acima das expectativas em relação ao número de dadores nestas áreas. Tal facto pode ser explicado pelo elevado número de população com idades compreendidas entre os 45 e os 65 anos de idade. Todavia, é nos hospitais que têm por área de abrangência municípios como Lisboa, Almada, Viana do Castelo, Torres Vedras, Castelo Branco e Leiria que se pode encontrar potencial para doação, uma vez que o número de dadores nestas áreas se encontra abaixo do que seria expectável, e, ademais, estes municípios são áreas com uma elevada população residente com idades compreendidas entre os 45 e os 65 anos. É de relevância levar a cabo ações de sensibilização nesses hospitais, fazendo chegar às pessoas mais informação relativamente ao tema.

Quanto aos doentes transplantados, é nas cidades pertencentes aos grandes centros urbanos da Área Metropolitana do Porto e da Área Metropolitana de Lisboa, seguindo-

se cidades como Coimbra, que se transplanta um maior número de doentes, o que em parte se explica pelo facto de serem os municípios com mais população. Porém, face à incidência de transplantados, ao considerar a população residente por município, o panorama altera-se visto que a incidência passa a ser elevada sobretudo no interior do país. Tal facto leva-nos a constatar uma resposta positiva às necessidades dos doentes transplantados residentes um pouco por todo o país, em especial no interior.

De acordo com os resultados da análise de potencial, municípios como Sintra, Coimbra, Valongo, Cascais e Montijo estão acima das expectativas no que diz respeito ao número de doentes transplantados. Já Lisboa, Santa Maria da Feira, Funchal e Almada encontram-se abaixo do que seria expectável no que se refere ao número de doentes transplantados nesses municípios. Estas são áreas onde se deve intervir de modo a prevenir e compreender o porquê do agravamento de certas patologias em determinados territórios.

Espacialmente, os fluxos de acesso de doentes transplantados, considerando o seu local de residência e a respetiva deslocação à unidade de transplantação, levaram a constatar que a maioria dos doentes renais transplantados nas unidades do Porto e de Lisboa residem próximos a estas. Pelo contrário, na unidade de transplantação de Coimbra pôde ser apurada uma maior dispersão geográfica face ao acesso ao centro de transplante, dado que esta unidade transplanta doentes originários um pouco de todo o país.

Quanto a doentes transplantados por fígado, a concentração de fluxos situa-se sobretudo na faixa litoral do continente, onde existe na maioria dos casos proximidade às unidades de transplantação. Contrariamente, no que concerne aos doentes transplantados por coração, cujo fluxo de acesso de um modo geral é disperso, pôde averiguar-se uma preferência pela unidade de transplantação de Coimbra no caso de doentes das regiões insulares.

O fluxo de doentes transplantados com pulmão é denso, dado que a maioria dos dadores tem por residência a região norte do continente. O Hospital de Santa Maria em

Lisboa é o único centro que realiza cirurgia de transplantação pulmonar, todavia, a maioria dos doentes a serem transplantados com pulmão têm como residência a região Norte, o que leva a que estes se tenham de deslocar obrigatoriamente à cidade de Lisboa.

Relativamente a doentes transplantados com pâncreas, o número de doentes é diminuto em comparação com as restantes tipologias, e a maioria dos doentes realiza a cirurgia na unidade mais próxima ao seu local de residência.

Para finalizar, importa referir que este estudo contou com algumas dificuldades, nomeadamente na obtenção e tratamento de dados, uma vez que grande parte da informação se encontrava em listagens em formato PDF ou disponíveis em papel. Este é um processo moroso atendendo ao volume de dados, pelo que é fulcral uma elevada concentração de modo a reduzir ao máximo possíveis erros humanos.

Existem duas possíveis formas de ajudar a reduzir as listas de espera de doentes para transplante. Uma hipótese passa pela identificação de dadores no território, e interligar estes às áreas de abrangência hospitalar, de modo a detetar potencial de doação nos hospitais. Uma segunda possibilidade passa pelo estudo de patologias associadas aos doentes que necessitam de um transplante. Como perspetivas futuras, resalto o estudo aprofundado dessas patologias no território ao longo dos anos, a sua incidência, assim como, de que modo estas se relacionam com as várias tipologias de doentes.

Referências Bibliográficas

- Acharya, U., Bhat, P., Iyengar, S., Rao, A. e Dua, S. (2003). *Classification of heart rate data using artificial neural network and fuzzy equivalence relation*. Pattern Recognition Society, vol. 36.
- Alves, R. (2016). *Aplicação de Modelos de Redes Neurais para Previsão de Consumos de Energia*. Dissertação de Mestrado em Engenharia Mecânica, Instituto Superior Técnico, UL, Lisboa.
- Anscombe, F. J. (1973). *Graphs in Statistical Analysis*. American Statistician. 27 (1): 17–21.
- Anselin, L. (1992). *Spatial Data Analysis with GIS: An Introduction to Application in the Social Sciences*. National Center for Geographic Information and Analysis. Technical Report. 3-54.
- Anselin, L. (1995). *Local indicators of spatial association-LISA*. Geographical Analysis. 27: 93-115.
- Anselin, L., Syabri I. e Kho Y. (2010). *Geodata: An Introduction to Spatial Data Analysis*. in Handbook of Applied Spatial Analysis Software Tools Methods and Applications. 73-87.
- Arbeit, D. (1992). *Geographic Information-Systems - a Guide to the Technology*. Journal of the American Planning Association, 58, (4): 531-532.
- Barcellos, C. (2008). *A geografia e o contexto dos problemas de saúde*. Saúde Movimento. Rio de Janeiro: 33-37.
- Batista, R., Vitorino, R., Gomes A., Oliveira, A., Ferreira, R., António, V. (2014). *As redes neuronais artificiais e o ensino da medicina*. Revista bras. de educação, vol. 38, Rio de Janeiro. doi: 10.1590/S0100-55022014000400017.
- Bickmore, D. P. e Tulloch, T. (1979). Medical Maps - some design suggestions for mapping health care and diseases statistics. Proceedings Auto - (2): 324 325.
- Blossom, C., Finkelstein L., Weihe G., and Bonnie B. (2011). *Applying GIS Methods to Public Health Research at Harvard University*. Journal of Map & Geography Libraries 7 (3) September: 349-376. doi: 10.1080/15420353.2011.599770.

- Box, G. E. P. (1979). *Robustness in the strategy of scientific model building*; in Launer, R. L.; Wilkinson, G. N., *Robustness in Statistics*, Academic Press, pp. 201–236.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software.
- Brunsdon, C. (2008). *Exploratory spatial data analysis (ESDA)*. in Encyclopedia of geographic information science, 136-140. Kemp K (Ed.). SAGE Publications. ISBN 978-1-4129-1313-3
- Buch, M. (2000). *Subsymbolic Techniques: Neural Networks*. Universidade de Kaiserslautern.
- Burrough, P. A. e McDonnell, R. A. (1998). *Principles of Geographical Information Systems*. Oxford University Press, New York. 333.
- Câmara, G., Monteiro, A., Fucks, S., Sá, M. (2004). *Spatial Analysis and GIS: A Primer*. Nacional Institute for Space Research (INPE), Brasil.
- Cerda, J. e Valdivia, G. (2007). *John Snow, la epidemia de cólera y el nacimiento de la epidemiología moderna*. Pontificia Universidad Católica de Chile. Facultad de Medicina, Departamento de Salud Pública. (24): 331-334.
- Cioca, M. e Cioca, L. (2010). Decision Support Systems used in Disaster Management. Decision Support Systems, Chiang S. Jao (Ed.) doi: 978-953-7619-64-0.
- Costa, R. (2017). *Aspectos jurídicos da transplantação de órgãos*. Dissertação de Mestrado em Direito, Universidade de Coimbra, 23:24. doi: hdl.handle.net/10316/84267
- Croner, M., Sperling, J. e Broome R. (1996). *Geographic information systems (GIS): New perspectives in understanding human health and environmental relationships*. Statistics in Medicine, (15): 17:18.
- De'ath, G. and Fabricius, K. E. (2000). *Classification and regression trees: a powerful yet simple technique for ecological data analysis*. Ecology, 81: 3178-3192.
- Demuth, H., Beale, M. e Hagan M. (2015). *Neural Network Toolbox TM User's Guide*. Natick, Massachusetts: The MathWorks, Inc.
- Direção Europeia da Qualidade dos Medicamentos e Cuidados de Saúde do Conselho da Europa (2016). *Guia para a qualidade e segurança dos órgãos para transplantação*. IPST, 5ª edição, Lisboa.

- Ebdon, D. (1985). *Statistics in geography*. Blackwell's Book Services, Oxford.
- Elzubair, R. (2009). *Aiding decision making in public health domain using WebGIS*. University Teknologi Malaysi.
- Encalada, L. (2016). *Big Data e Redes Sociais na Análise Espacial de Padrões Turísticos em Contexto Urbano*. O caso de Lisboa. Dissertação de Mestrado, 142:143.
- Fonseca, L. (2012). *Contributo dos SIG para a definição de áreas geográficas para a distribuição de enfermeiros de família: Estudo de caso na USF Nova Via do ACES Espinho/Gaia*. Faculdade de Letras, Universidade do Porto.
- Fradelos, C., Papathanasiou, I., Mitsi, D., Tsaras, K., Kleisaris, C. e Kourkouta, L. (2014). *Health based geographic information systems (GIS) and their applications*. Acta Inform. Medical, (22): 402–405. doi: 10.5455/aim.2014.22.402-405.
- Freeman, J. A. e Skapura, D. M. (1992). Backpropagation. *Neural Networks Algorithm Applications and Programming Techniques*. 89-125.
- Gallant, S. (1993). *Neural Network Learning and Expert Systems*. A Bradford Book, Cambridge, Massachusetts, London.
- Gallo, J. e Ertur, C. (2000). *Exploratory Spatial data analysis of the distribution of regional per capita GDP in Europe, 1980-1995*. Pôle d' Économie et de Gestion. 2-9.
- Gao, S. (2010). *Advanced Health Information Sharing with Web-Based GIS*. Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada.
- Gatrell, A. e Bailey T. (1995). *Can GIS be made to sing and dance to an epidemiological tune*. Presented at the International Symposium on Computer Mapping and Environmental Health, Tampa, FL, February.
- Getis, A. (1999). *Spatial statistics, in Geographical Information Systems*. Principles, Techniques, Management and Applications, (1): 283-292.
- Goldberg, D. W., Wilson, J. P. e Knoblock, C. A. (2007). *From Text to Geographic Coordinates: The Current State of Geocoding*. URISA J, (19): 33-46.
- Goodchild, M. (2003). *Geographic Information Science and Systems for Environmental Management*. Annual Review of Environment and Resources, (28): 493-519. doi: 10.1146/annurev.energy.28.050302.105521.

- Gorry, G. A. e Morton, M. S. (1971). *A Framework for Management Information Systems*. Sloan Management Review 13, (1): 55-70.
- Hair, J. F. (1995). *Multivariate Data Analysis*. Seventh Edition.
- Hanchette, C. L. (2003). *Public Health Informatics and Information Systems*. em: O'Carroll, P.W., Ripp, L.H., Yasnoff, W.A., Ward, M.E., Martin, E.L. (Eds.), Springer, New York, NY: 431-466.
- Harries, K. (1999). *Mapping Crime: Principle and Practice*. Washington D.C. Newsletter Transplant International figures on donation and transplantation 2016 – vol. 22.
- Haykin, S. (1999). *Neural Networks – A Comprehensive Foundation*. Prattice-Hall, New Jersey, 2nd edition.
- Heidari, M. e Omid, M. (2011). *Appllication of Artificial Neural Network for the Modeling Benefit to Cost*. Journal of Applied Science, Engineering and technology, University of Tehran, Iran.
- Hill, L. (2006). *Georeferencing: The Geographic Associations of Information*. The Mit Press, 260. doi:262-08354-X.
- Keenan, P. B. (2003). *Spatial Decision Support Systems*. In *Decision making support systems: Achievements, trends and challenges for the new decade*, ed. M. Mora, G. Forgionne, and J. N. D. Gupta, 28-39. Hershey, PA: Idea Group Publishing.
- Kemp, K. K. (2008). *Encyclopedia of Geographic Information Science*. SAGE. doi: 10.4135/9781412953962
- Kohonen, T. (2014). *MATLAB Implementations and Applications of the Self-Organizing Map*. Alto University, School of Science, Helsinki, Finland.
- Lee, J. e Wong, D. (2001). *Statistical analysis with ArcView GIS*. John Wiley & Sons.
- Lippmann, R. (1987). *An introduction to computing with neural nets*. In IEEE ASSP Magazine. doi: 10.1145/44571.44572.
- Longley, P. A., Goodchild, M. F., Maguire, D. J. e Rhind D. W. (2011). *Geographical Information Systems and Science*. doi: 978-0-470-72144-5.
- Longley, P. A., Goodchild, M. F., Maguire, D. J. e Rhind D. W. (2005). *Geographical Information Systems and Science*. John Wiley & Sons Ltd, 2^a nd edition, England.

- Lumpkin, J. R. (2003). *History and significance of information systems and public health*. In P.W. O'Carroll, Y.A. Yasnoff, M.E. Ward, L.H. Ripp, and E.L. Martin (Ed.), *Public Health Informatics* (pp. 16–38). New York, NY: Springer.
- Malczewski J. (1997). *Spatial Decision Support Systems*. NCGIA Core Curriculum in GIScience.
- Miedes, A. (2008). *La Actitud e Intención de la Donación de Órganos em la Población Espanola: Análisis Mediante Regresión Logística Multinivel*.
- Mileu, N. (2016). *Sistema de Apoio à Decisão na Gestão do Risco à Escala Municipal*. Dissertação de Doutoramento, Instituto de Geografia e Ordenamento do Território, Lisboa.
- Mitchell, A. (2005). *ESRI Guide to GIS Analysis - Volume 2: Spatial Measurements & Statistics*. ESRI Press, RedLands, California.
- Mitra, S. (2009). *Spatial Autocorrelation and Bayesian Spatial Statistical Method for Analyzing Fatal-and Injury-Crash-Prone Intersections*. Presented at 88th Annual Meeting of the Transportation Research Board.
- Moore, D. e Carpenter, T. (1999). *Spatial analytical methods and geographic information systems: use in health research and epidemiology*. *Epidemiol Rev.* Volume 21 nº2: 143–161. doi: 10.1.1.114.2605.
- Nogueira, H. (2007). *Os Lugares e a Saúde - Uma abordagem da Geografia às variações em saúde na Área Metropolitana de Lisboa*. Dissertação de Doutoramento, Faculdade de Letras da Universidade de Coimbra.
- Nogueira, H. e Remoaldo, P. (2010). *Olhares geográficos sobre a Saúde*. Colibri. doi: 978-972-772-995-1: 31-33.
- Nossa, P (2005). *Abordagem Geográfica da Oferta/Procura dos Cuidados de Saúde, o caso da Sida*. Tese de Doutoramento em Geografia (Geografia Humana), Universidade do Minho.
- Oliveira, P. (2011). *Transplante Renal: Desigualdades no Acesso*. Dissertação em Gestão da Saúde, Universidade Nova de Lisboa, Lisboa.
- Ozernoy, M., Smith, R. e Sicherman, A. (1981). *Evaluating Computerized Geographic Information-Systems Using Decision-Analysis*. *Interfaces*, 11(5): 92-100.

- Paiva J., Fernandes A., Granja, C., Esteves, F., Ribeiro J., Nóbrega J., Vaz, J., Coutinho, P. (2017). *Rede Nacional de Especialidade Hospitalar e de Referência – Medicina Intensiva*. República Portuguesa, Saúde.
- Paradis, E. (2018). *Moran's Autocorrelation Coefficient in Comparative Methods*. 1-9.
- Power, D. J. (2007). *A Brief History of Decision Support System*. DSSResources.COM, World Wide Web.
- Ricketts, C. (2003). *Geographic information systems and public health*. Annual Review of Public Health 24: 1-6.
- Riner, M. E., Cunningham, C. e Johnson, A. (2004). *Public health education and practice using geographic information system technology*. Public Health Nursing, 21(1): 57-65.
- Ripley, B. D. (1996). *Pattern recognition and neural networks*. Cambridge: Cambridge University Press.
- Rocha, J., (2012). *Sistemas Complexos, Modelação e Geosimulação da Evolução de Padrões de Uso e Ocupação do Solo*. Diss. de Doutoramento em Geografia.
- Rocha, J., Tenedório, J., Encarnação, S, e Pontes, S. (2014). *Integração de Autómatos Celulares e Redes Neurais em Ambiente SIG para Previsão de Alterações nos Padrões de Uso e Ocupação do Solo*. CEG, U.L e CEGPR, U.N.L., Lisboa.
- Roquette, R., Nunes, B. e Painho, M. (2018). *The relevance of spatial aggregation level and of applied methods in the analysis of geographical distribution of cancer mortality in mainland Portugal (2009-2013)*. Population Health Metrics, 16:6. doi: org/10.1186/s12963-018-0164-6
- Rubio, N. e Alba, L. (2000). *Quies ser donante? Haz que todos lo sepan. Donació transplantaments*. Revista Viure en Salut, nº47:7.
- Rumelhart, D. E., Hilton, G. E. e Williams, R. J. (1986). *Learning internal representations by error propagation*. Parallel Distributed Processing, Vol. 1, Mit Press: 318-362.
- Rushton, G. (2003). *Public health, GIS, and spatial analytic tools*. Annual Review of Public Health. 24 (1): 43-56.
- Russell, S. J. e Norvig P. (1995). *Artificial Intelligence: a modern approach*. New Jersey: Prentice-Hall Inc.
- Saldiva, D. e Alvarenga, G. (2003). *Correlação e Regressão*. Tempo 1–8.

- Santana, P. (2014). *Introdução à Geografia da Saúde: Território, Saúde e Bem-Estar*. Coimbra University Press: 13-20.
- Santos, A., Seixas, J., Pereira B., e Medronho R. (2005). *Using Artificial Neural Networks and Logistic Regression in the Prediction of Hepatitis A*. Rev. bras. Epidemiol. Vol.8 nº 2: 119:120. doi: 10.1590/S1415-790X2005000200004
- Scott, L. e Janikas, M. (2010). *Spatial statistics in ArcGIS, in Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*. 27 - 40.
- Seaman, V. (1798). *Inquiry into the cause of the prevalence of yellow fever in New York*. Medical Repository 1: 314.
- Shekar, S. e Xiong, H. (2008). *Encyclopedia of GIS, Control*. Springer New York.
- Shimazono, Y. (2007). *The state of the international organ trade: a provisional picture based on integration of available information*. Institute of Social and Cultural Anthropology, University of Oxford, Oxford, England.
- Siffel, C., Strickland, M. J., Gardner, B. R., Kirby R. S. e Correa, A. (2006). *Role of geographic information systems in birth defects surveillance and research*. Birth Defects Research Part a-Clinical and Molecular Teratology, 76 (11): 825-833.
- Snow, J. (1855). *On the mode of communication of cholera*. John Churchill, London.
- Sugumaran, R. e DeGroote, J. (2011). *Spatial Decision Support Systems: Principles and Practices*. CRC Press, Taylor & Francis Group.
- Sugumaran, V. e Sugumaran, R. (2005). *Web-based Spatial Decision Support Systems (WebSDSS): evolution, architecture, and challenges*. Third Annual SIGDSS Pre-ICIC Workshop Designing Complex Decision Support: Discovery and Presentation of Information and Knowledge, Las Vegas.
- Taylor, P. J. e Overton, M. (1991). *Further Thoughts on Geography and GIS - a Preemptive Strike*. Environment and Planning A, 23(8): 1087-1090.
- Tobler, W. R. (1970). *A Computer Movie Simulation Urban Growth in Detroit Region*. Econ. Geography. 46: 234-240. doi: 10.1126/science.11.277.620.
- Venables, W. N. and Ripley, B. D. (2002). *S Programming*. 4th edition. New York: Springer-Verlag
- Waller, L. e Gotway C. (2004) *Applied spatial statistics for public health data*. John Wiley & Sons, Inc., Hoboken, New Jersey. ISBN 0-471-38771-1

- Zambon, M., Lawrence, R., Bunn, A., and Powell, S. (2006). *Effect of alternative splitting rules on image processing using classification tree analysis*. Photogrammetric Engineering and Remote Sensing, 72(1): 25–30.
- Zimmerman, D. L., Fang, X. M. e Mazumdar S. (2008). *Spatial clustering of the failure to geocode and its implications for the detection of disease clustering*. Statistics in medicine, 27(21): 4254-4266.

Legislação Consultada

Artigo 1º da Portaria nº 76/2014, de 21 de março e de acordo com a Lei nº 36/2013, de 12 de junho – Diário da República, 1ª série, nº 57 de 21 de março de 2014. Consultado a 16/05/2018.

Artigo 10º da Portaria nº 22/2007, de 29 de junho – Diário da República, 1ª série, nº 124 de 29 de junho de 2007. Consultado a 12/06/2018.

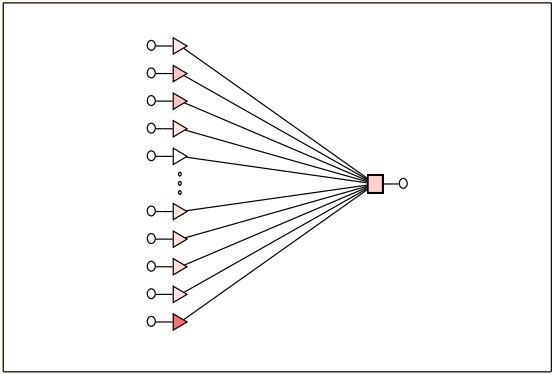
Artigo da Portaria n.º 357/2008, de 09 de maio – Diário da República, 1ª série, nº 90 de 9 de maio de 2008. Consultado a 18/05/2018.

Despacho nº 5480/2017 - Diário da República, 2.ª série n.º 120 de 23 de junho de 2017. Consultado a 04/05/2018.

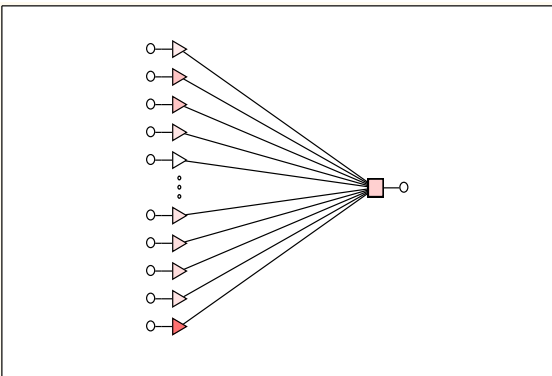
Despacho nº 14341/2013, 29 de outubro de 2013 – Diário da República, 2ª série nº 215 de 6 de novembro de 2013. Ministério da Saúde - O Secretário de Estado Adjunto do Ministro da Saúde, Fernando Serra Leal da Costa. Consultado a 16/06/2018.

Anexos

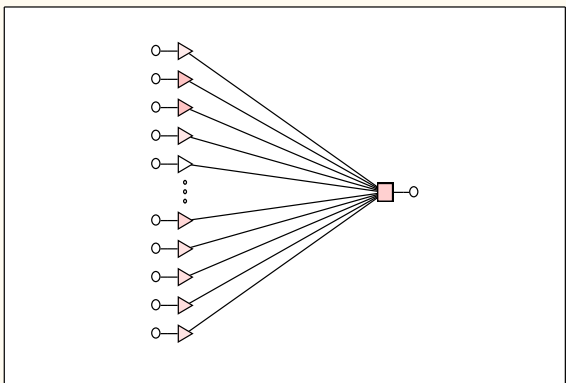
Anexo 1: Arquitetura dos modelos de RNA testados no estudo – Dadores de órgãos falecidos.



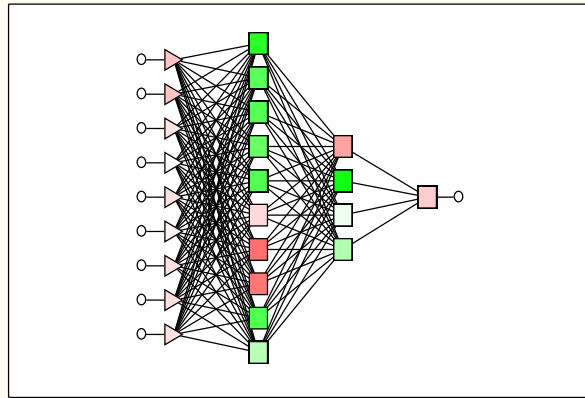
Linear 1



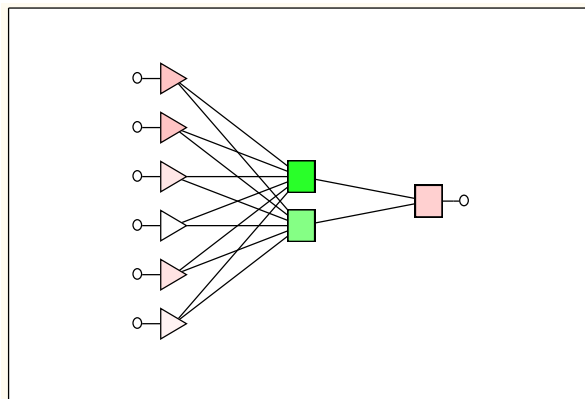
Linear 2



Linear 3

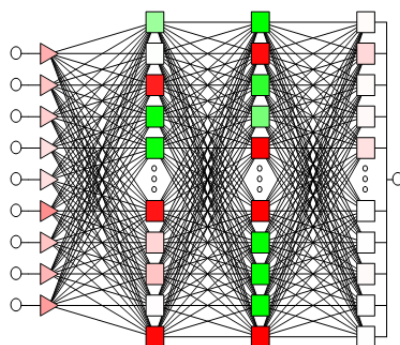


MLP 4

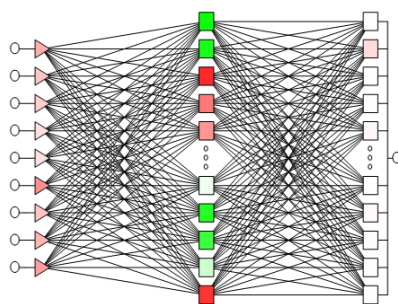


MLP 5

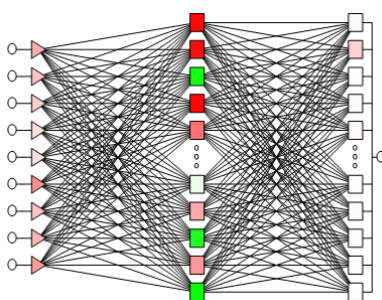
Anexo 2: Arquitetura dos modelos de RNA testados no estudo – Doentes transplantados.



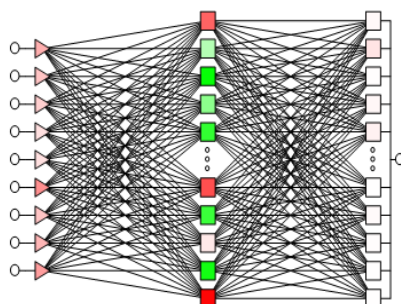
MLP 1



MLP 2



MLP 3



MLP 4

